AD_____

Award Number: W81XWH-09-1-0488

TITLE: Systematic Search for Gene-Gene Interaction Effect on Prostate Cancer Risk

PRINCIPAL INVESTIGATOR: Jielin Sun, Ph.D.

CONTRACTING ORGANIZATION: Wake Forest Universiyt Health Sciences
Winston Salem, NC 27157

REPORT DATE: July 2013

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

Approved for public release; distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
July 2013	FINAL	01 July 2009 – 30 June 2013

4. TITLE AND SUBTITLE
Systematic Search for Gene-Gene Interaction Effect on Prostate Cancer Risk

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-09-1-0488

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jielin Sun; Jianfeng Xu, Siqun Lilly Zheng

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Wake Forest University Health Sciences

Winston Salem, NC 27157

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Prostate cancer (PCa) is the most common cancer and the second leading cause of cancer death among men in the United States. Considering that PCa development requires the coordination of many genes, it is expected that a simultaneous evaluation of multiple genetic variants can improve the statistical power to detect additional PCa risk variants. Recent improvements in analytical methods and computation make it feasible to search for gene-gene interaction of SNPs in the genome. We hypothesized that multiple sequence variants in the genome may interact to increase PCa risk. These variants may or may not have known main effect on PCa risk and can be better detected by systematically evaluating gene-gene interactions for SNPs in the genome. We utilized data from an existing GWAS of a large NCI Cancer Genetic Markers of Susceptibility (CGEMS) study to systematically discover genes that interacted with known PCa risk variants in the genome. We also evaluated the genes that interacted with known PCa risk variants in another two independent populations, including a population based PCa case-control study from Sweden (CAPS) and a PCa patient population from Johns Hopkins Hospital (JHH). In addition, we performed an exhaustive search for pair-wide SNP-SNP interactions without main effect in the JHH and CGEMS populations using a novel statistical approach of Boolean Operation-based Screening and Testing (BOOST). We identified thirty-five pairs of SNPs that significantly interacted with the thirty-two known risk variants on PCa risk at a P-value of 1E-05 in the combined analysis of three populations. The most significant interaction detected was between rs12418451 in *MYEOV* and rs784411 in *CEP152*, with a $P_{interaction}$ of 1.15E-07 in the meta-analysis. We also emphasized two pairs of interactions with potential biological implication, including an interaction between rs7127900 near *IGF2/IGF2AS* and rs12628051 in *TNRC6B*, with a P interaction of 3.39E-06; and an interaction between rs7679763 near *TET2* and rs290258 in *SYK*, with a P interaction of 1.49E-06. In addition, BOOST analysis revealed an additional 16 pairs of gene-gene interactions without main effects that confer risk to prostate cancer in both CGEMS and JHH populations. Those results show statistical evidence for novel loci interacting with or without known risk-associated SNPs to modify PCa risk. The interacting loci identified provide hints on the underlying molecular mechanism of the associations with PCa risk for the known risk-associated SNPs. Our study also represents the first step towards obtaining further biological insight into the high-dimensional etiology of prostate cancer.

15. SUBJECT TERMS
Risk SNPs, Interaction

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON USAMRMC**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UU	31	**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Page

INTRODUCTION

Prostate cancer (PCa) is the leading cancer among men in the United States, and is a disease with strong genetic susceptibility. The genetic susceptibility is due to the inheritance of multiple sequence variants, majorly in the form of single nucleotide polymorphisms (SNPs). Most current genetic studies focus only on the single SNP association studies. In contrast, few studies have explored the role of interactions of multiple SNPs with PCa risk, due to limited statistical approaches available to study interactions in a genome-wide level. In fact, gene-gene interaction is the norm rather than exception for complex diseases such as PCa. Inference from tumorigenesis and results from genetic modeling studies suggest that multiple susceptibility genes, either additively or multiplicatively, determine individual risk to PCa. The importance of gene-gene interaction is also supported by empirical evidence from model organisms and human studies. Results from simulation studies suggest that simultaneous evaluation of multiple genetic variants can improve the statistical power to detect additional PCa risk variants and can be more fruitful than traditional approaches that exclusively focus on main effects34. It is expected that additional PCa risk variants will be identified using gene-gene interaction analyses.

In this DOD funded proposal, we propose to 1) identify SNPs in the genome that interact to have stronger effects on PCa risk in the CGEMS GWAS data, 2) confirm the gene-gene interaction effect on PCa risk identified from the CGEMS study in 1,000 PCa cases and 1,000 controls in CAPS, 3) further confirm the gene-gene interaction effects on PCa risk for pairs of SNPs implicated in Aim 2 among the remaining 1,893 cases and 781 controls in CAPS, and 4) fine map the genomic regions where SNPs have been confirmed to have a strong gene-gene interaction effect on PCa risk among all 2,893 cases and 1,781 controls in CAPS.

BODY

Approved Statement of Work:

STATEMENT OF WORK

We will take advantage of two existing large study populations to test this hypothesis. The first study population is the NCI Cancer Genetic Markers of Susceptibility (CGEMS) study, where genome-wide data are available for 1,172 PCa case patients and 1,157 control subjects of European Americans descent. The second study population is our large population-based case-control study from Sweden (CAPS), with 2,893 PCa patients and 1,781 control subjects. These two large study populations are necessary to have sufficient statistical power to detect true interactions in the genome and to remove false positives using a multiple-stage study design.

Aim 1. Identify SNPs in the genome that interact to have stronger effects on PCa risk in the CGEMS GWAS data.

Step by step method and expected results

1) **(Months 1)** Preparation of the study, including IRB and other logistic issues.
2) **(Months 2)** Pre-association analysis and imputation of all the SNPs in the genome.
3) **(Months 3-12)** Logistic regression analysis to identify pairs of SNPs that interact to have stronger effects on PCa risk.

4) **(Months 3-12)** Bayesian epistasis association mapping (BEAM) to identify pairs of SNPs that interact to have stronger effects on PCa risk.

Outcome and deliverables

We will identify pairs of SNPs in the genome whose interaction terms reach $P < 0.001$. Based on the results of genome-wide searches for SNPs that interact with rs1447295 at 8q24, we expect to identify ~450 independent SNPs that interact with each of the twelve known PCa risk SNPs. In addition, we expect to identify additional SNPs that have stronger two-way or three-way interaction effects on PCa risk from the BEAM analysis. These SNPs, estimated to number around ~5,000, will be selected for further confirmation in Aim 2.

Aim 2. Confirm gene-gene interaction effects on PCa risk identified from the CGEMS study in 1,000 PCa cases and 1,000 controls in CAPS

Step by step method and expected results

1) **(Months 6-16)** Genotype ~5,000 SNPs among 1,000 cases and 1,000 controls in CAPS using iSELECT of Illumina.
2) **(Months 17-20)** Analyze data using logistic regression and BEAM methods to remove false positives identified in Aim 1 and obtain a smaller subset of SNPs that most likely represent true interaction for further confirmation in larger samples.

Outcome and deliverables

Most of the false positive interactions will be removed from this aim. We expect 125 pairs of SNPs in this stage will have $P < 0.05$ and have the same direction of interaction effect as in the CGEMS data [(5,000 x 0.05)/2]. These SNPs will be selected for further confirmation in Aim 3. The actual number of SNPs may be higher if there are more true interaction effects in the genome.

Aim 3. Further confirm the gene-gene interaction effects on PCa risk for pairs of SNPs implicated in Aim 2 among the remaining 1,893 cases and 781 controls in CAPS

Step by step method and expected results

1) **(Months 21-24)** Genotype 125 SNPs among 1,893 cases and 781 controls in CAPS using iPLEX of Sequenom.
2) **(Months 25-28)** Analyze data using logistic regression and BEAM methods to remove false positives identified in Aim 1 and obtain a smaller subset of SNPs that most likely represents true interaction for further confirmation in larger samples.

Outcome and deliverables

We will use a stringent genome-wide significance level from the joint analysis ($P < 10^{-8}$) to declare significant interaction. In fact, considering that 12 genome-wide associations were performed, it is more conservative to use the cutoff of $P < 10^{-9}$. The actual number of SNPs whose interactions meet the criteria depends on the number of true interactions in the genome with the OR detectable in our study.

Aim 4. Fine map the genomic regions where SNPs have been confirmed to have strong gene-gene interaction effect on PCa risk among all 2,893 cases and 1,781 controls in CAPS

Step by step method and expected results

1) **(Months 29-30)** Functional SNPs and tagging SNPs will be selected in the genomic region for each of the SNPs implicated in Aim 3.
2) **(Months 29-30)** Selected SNPs will be genotyped in 2,893 cases and 1,781 controls of CAPS using iPLEX of Sequenom.
3) **(Months 31-36)** Fine mapping data analysis will be performed to identify the most strongly associated SNPs (interaction effect) in each of the regions implicated in Aim 3.

Additional statistical analysis on gene-gene interaction (month 37- 48, non-cost extension). We received approval from DOD to perform additional gene-gene interaction analysis based on novel statistical approaches developed recently, which were not originally proposed in the study.

Detailed report

Study design modification. In our initial report, we proposed to conduct a genome-wide search in the CGEMS population and follow the top hits in an additional two study populations (CAPS and JHH). During year 3, we were able to obtain access to the GWAS data for the CAPS and JHH populations. Therefore, we also conducted a genome-wide search for SNPs that interact with the 32 risk SNPs using CAPS and JHH populations. We also performed a meta-analysis and a fine-mapping study based on the GWAS data of the three populations. Compared with our original study design, the new design greatly improved our statistical power to detect SNP-SNP interactions. For example, we were only able to detect a relatively large effect (OR>1.7) based on our initial design of using 1,176 cases and 1,101 controls from the CGEMS study. However, we were able to detect a modest to large effect of interaction (OR>1.3) using a total of 4,723 PCa cases and 4,792 controls based on three GWAS populations.

Study populations. The first population was obtained from Stage 1 of the National Cancer Institute Cancer Genetic Markers of Susceptibility (CGEMS) study. It included 1,176 PCa cases and 1,101 control subjects, selected from the Prostate, Lung, Colon and Ovarian (PLCO) Cancer Screening Trial. The genotype and phenotype data of the study are publicly available and our use of the data was approved by CGEMS.

The second GWAS population included 1,583 prostate cancer patients and 519 control subjects that matched the age distribution of case subjects from CAPS, a population-based PCa case-control study from Sweden (CAPS). Briefly, the CAPS population was recruited from four regional cancer registries in Sweden and diagnosed between July 2001 and October 2003. The clinical characteristics of these patients are presented in **Supplementary Table 1**.

The third population was from a Johns Hopkins Hospital (JHH) PCa GWAS which included 1,964 PCa cases and 3,172 control subjects. The cases are Caucasian PCa patients who underwent radical prostatectomy for the treatment of PCa at JHH from January 1, 1999, through December 31, 2008 [1]. The clinical characteristics of these patients are presented in **Supplementary Table 2**. The control subjects for this population were an independent group of Caucasian individuals from the Illumina iControlDB (iControls) dataset (https://www.illumina.com/science/icontrodb.ilmn).

Genotype data, imputation, and quality control. GWAS of the CAPS population was performed using Affymetrix 5.0 chip. GWAS of the JHH case population was performed using the Illumina 610K chip. GWAS of the iControls population was performed using Illumina Hap300 and Hap550 Chips. GWAS of the CGEMS population was performed using HumanHap300 and HumanHap240 assays from Illumina Corp.

For each GWAS population, we imputed all the known SNPs that are catalogued in HapMap Phase II (www.hapmap.org) using the IMPUTE computer program [2] with a posterior probability of 0.9 as a threshold to call genotypes. Individuals with a call rate below 0.95 were removed from GWAS analysis. The following quality control criteria were used to filter SNPs: MAF < 0.01, HWE < 0.001 and call rate < 0.95.

Prostate cancer known risk SNPs identified from GWAS. The 33 PCa known risk-associated SNPs were discovered by GWAS and the following fine-mapping studies, with *P*- values equal or smaller than of 10E-7 [3-17]. The detailed information for the 33 risk SNPs are presented **in Table 1**. The SNP rs16901979 was not evaluated in the following interaction analysis due to the unavailability of imputation of this SNP since it was not catalogued in the HapMap database.

Statistical analysis.

Part A: Identify SNPs that interact with 32 known PCa risk variants.

Multiplicative interactions between each one of the 32 known PCa risk variants and each SNP in the genome were systematically tested by including both SNPs and an interaction term (product of two SNPs), as implemented in the computer program PLINK[18]. Ancestral proportions obtained based on EIGENSOFT software [19]were included as covariates to minimize the impact of potential population stratification in the JHH population. An additive genetic model was used, where the genotypes were coded as 0,1, and 2 and each SNP was treated as a continuous variable. The interaction term was tested using a Wald test, with degree of freedom of 1. A meta-analysis of the interaction term for the three study populations was performed using the method developed by Manning et al [20] . Briefly, the meta-OR (OR_M) of the interaction term across the three populations was estimated using an inverse-variance weighted meta-analysis where

$\ln(OR_M) = \sum_{i=1}^{3}(w_i \ln(OR_i)/\sum_{i=1}^{3} w_i),\ w_i = 1/\mathrm{var}(\ln(OR_i))$ and

$se(\ln(OR_M)) = [\sum_{i=1}^{3}(1/\mathrm{var}(\ln(OR_i))]^{-1/2}$.(20)

Part B: Identify SNP-SNP interaction that without main effects using a novel statistical method of BOOST

We implemented the Boolean Operation-based Screening and Testing (BOOST) approach to identify gene-gene interaction on the genome-wide scale [21]. BOOST utilizes log-linear model to evaluate a two-locus interaction, which is much faster than the logistic model since it does not require iteration. It has been shown that using Log-linear model is equivalent with the logistic model [21]. Since there are three genotypes for each SNP, a 3x3x2 contingency table was constructed to test the interaction effect of SNP pair (X_P,X_Q). The log-linear models for the homogeneous association model (M_H) with main effect terms and the saturated model (M_S) with interaction term are as following:

$$M_H : \log\mu_{ijk} = \lambda + \lambda_i^{X_p} + \lambda_j^{X_q} + \lambda_k^{Y} + \lambda_{ij}^{X_pX_q} + \lambda_{ik}^{X_pY} + \lambda_{jk}^{X_qY}$$

$$M_S : \log\mu_{ijk} = \lambda + \lambda_i^{X_p} + \lambda_j^{X_q} + \lambda_k^{Y} + \lambda_{ij}^{X_pX_q} + \lambda_{ik}^{X_pY} + \lambda_{jk}^{X_qY} + \lambda_{ijk}^{X_pX_qY}$$

where μ_{ijk} is the mean of the cell count of the contingency table corresponding to SNP *Xp* and *Xq* and disease status *Y*, λ is the overall mean of the natural log of the expected frequencies, and λs are the corresponding effects which SNPs and disease status have on the cell frequencies. Here $\lambda_{ijk}^{X_p X_q Y}$ measures the gene-gene interaction and is the term of interest. If we use L_H and L_S to denote the log-likelihood of M_H and *Ms*. The significance of interaction term is tested by the difference of the maximum likelihood estimations: $L_S - L_H$.

A two-stage approach was then implemented to improve the computation efficiency while keeping the interactions in check. In the screening stage, BOOST transformed the genotype data into Boolean values and stored it into a contingency table. The difference of maximum likelihood estimation (MLE) of these two models was used to test interaction significance of a SNP pair. To overcome the computation difficulty that there is no close form estimate of MLE, BOOST approximates the likelihood ratio statistics *L* for the homogeneous association model using Kirkwood superposition approximation (KSA). Because $L_S - L_{KSA}$ is an upper bound of $L_S - L_H$, this method filters out most non-significant interactions but also guarantees survival of significant ones. A SNP pair would be removed from consideration if the difference of MLE between these two models ($2(L_S - L_{KSA})$) is less than a specified threshold *t*. In our analysis, we set *t* to be 43 for CGEMS and 9.49 for the confirmation population (JHH), respectively. This corresponds to an unadjusted $P_{\text{interaction}}$ of 1E-08, and a $P_{\text{interaction}}$ of 0.05, respectively. In the testing stage, likelihood ratio test statistics $2(L_S - L_H)$ was used to test the interaction of remaining SNP pairs. The test statistic is evaluated by a χ^2 test with four degrees of freedom. The $P_{\text{interaction}}$ can be further adjusted by Bonferroni correction, if needed. Ancestral proportions obtained based on EIGENSOFT software [19] were included as covariates to minimize the impact of potential population stratification in the JHH population. No age adjustment was performed due to the incomplete information in the iControls.

Results

Part A: Identify SNPs that interact with 32 known PCa risk variants.

After imputation and applying quality control (QC) criteria, 1,314,700, 1,646,196, and 1,757,946 SNPs remained for CAPS, JHH, and CGEMS studies, respectively. A total of 1,117,531 common SNPs for those three populations were used in the interaction analysis.

We examined the inflation factor and the quantile-quantile plots for interaction tests in the combined analysis of three populations. No systematic bias was observed, as the inflation factors for the 32 GWAS scans for SNP-SNP interactions ranged from 0.98 to 1.03 (Supplementary Table 3).

Table 1. Reported SNPs associated with Prostate Cancer

CHR	SNPs	Cytogenetic bands	Position	Known genes	m/M allele	Risk allele
2	rs1465618	2p21	43,407,453	*THADA*	A/G	A
2	rs721048	2p15	62,985,235	*EHBP1*	A/G	A
2	rs12621278	2q31.1	173,019,799	*ITGA6*	G/A	A

3	rs2660753	3p12	87,193,364	--	T/C	T
3	rs10934853	3q21.3	129,521,063	*EEFSEC*	A/C	A
4	rs17021918	4q22.3	95,781,900	*PDLIM5*	T/C	C
4	rs7679673	4q24	106,280,983	*TET2*	A/C	C
6	rs9364554	6q25	160,753,654	*SLC22A3*	T/C	T
7	rs10486567	7p15	27,943,088	*JAZF1*	A/G	G
7	rs6465657	7q21	97,654,263	*LMTK2*	T/C	C
8	rs2928679	8p21.2	23,494,920	*SLC25A37*	A/G	A
8	rs1512268	8p21.2	23,582,408	*NKX3.1*	T/C	T
8	rs10086908	8q24 (5)	128,081,119	--	C/T	T
8	rs16901979	8q24 (2)	128,194,098	--	A/C	A
8	rs16902094	8q24.21	128,389,528	--	N/A	G
8	rs620861	8q24 (4)	128,404,855	--	A/G	G
8	rs6983267	8q24 (3)	128,482,487	--	G/T	G
8	rs1447295	8q24 (1)	128,554,220	--	A/C	A
9	rs1571801	9q33	123,467,194	*DAB2IC*	G/A	A
10	rs10993994	10q11	51,219,502	*MSMB*	T/C	T
10	rs4962416	10q26	126,686,862	*CTBP2*	C/T	C
11	rs7127900	11p15.5	2,190,150	*IGF2, IGF2AS, INS, TH*	G/A	A
11	rs12418451	11q13 (2)	68,691,995	--	A/G	A
11	rs10896449	11q13 (1)	68,751,243	*MYEOV*	A/G	G
17	rs11649743	17q12 (2)	33,149,092	*HNF1B*	A/G	G
17	rs4430796	17q12 (1)	33,172,153	*HNF1B*	A/G	A
17	rs1859962	17q24.3	66,620,348	--	G/T	G
19	rs8102476	19q13.2	43,427,453	*PPP1R14A*	T/C	C
19	rs887391	19q13	46,677,464	--	C/T	T
19	rs2735839	19q13	56,056,435	*KLK3*	A/G	G
22	rs9623117	New 22q13	38,782,065	*TNRC6B*	C/T	C
22	rs5759167	New 22q13.2	41,830,156	TTLL1, BIK, MCAT, PACSIN2	T/G	G
23	rs5945619	Xp11	51,258,412	NUDT10, NUDT11, LOC340602	C/T	C

Abbreviation: Chr, chromosome; BP: Base pair position is based on NCBI build 36. m/M denotes minor allele/ major allele.

The results for the top ranked SNPs that interacted with each of the 32 known PCa-risk SNPs ($P_{interaction}$ <1.0E-05 in the meta-analysis) were presented in **Supplementary Table 4**. For SNPs in linkage disequilibrium (LD) (as defined by r2>0.5), only the one with the smallest *P* -value based on meta-analysis was included in the Supplementary Table 4. We then further examined the interaction effects for the top ranked SNPs ($P_{interaction}$ <1E-05*)* in each of the three populations. SNPs that significantly interacted with the 32 SNPs in all three populations at a nominal $P_{interaction}$ of 0.05 were presented in Table 2. No SNP-SNP interaction reached a genome-wide significant level of 1.5E-09 (0.05/(1e+6*32). The most significant interaction was observed between rs12418451 in the *MYEOV* gene region and rs784411 in the intron of *CEP152*, with a $P_{interaction}$ of 1.15E-07 ($OR_{interaction}$= 1.42; 95% CI: 1.25-1.61) in the meta-analysis. This interaction pair was significant in all three populations and the effects of the interaction were in the same direction ($P_{interaction}$ = 0.008, $OR_{interaction}$ =1.55 (95% CI: 1.12-2.16) for CAPS; $P_{interaction}$ = 0.005, $OR_{interaction}$ = 1.34 (95% CI: 1.14-1.58) for JHH; and $P_{interaction}$ = 0.001, $OR_{interaction}$ = 1.53 (95% CI = 1.18-1.99) in CGEMS, respectively) (Table 2).

Among the other 34 pairs of interactions that were significant at a $P_{interaction}$ cutoff of 1E-05 in the meta-analysis, two pairs were noteworthy to be emphasized when considering possible biological function. The first pair involved an interaction between rs7127900 at *IGF2/IGF2AS* region and rs12628051 in the intron of *TNRC6B*, with a $P_{interaction}$ of 3.39E-06 ($OR_{interaction}$ = 1.30, 95% CI = 1.17-1.46) (Table 2). The interaction was significant in all three populations and the effects of the interaction were in the same direction ($P_{interaction}$ = 0.002, $OR_{interaction}$= 1.50, 95 % CI = 1.16-1.93 in CAPS; $P_{interaction}$ = 0.006, $OR_{interaction}$ = 1.24, 95 % CI = 1.06-1.44 in JHH; $P_{interaction}$ = 0.014, $OR_{interaction}$ = 1.32, 95 % CI = 1.06-1.65 in CGEMS). The 2nd pair of interaction was between rs7679763 in *TET2* gene region and rs290258 in the promoter region of SYK, with a $P_{interaction}$ of 1.49E-06 (OR = 0.75, 95% CI = 0.67-0.84) (Table 2). Similarly, the interaction effect was consistently observed in all three populations with the same direction of interaction effect ($P_{interaction}$ = 0.002, $OR_{interaction}$ = 0.66, 95 % CI = 0.51-0.86 in CAPS; $P_{interaction}$ = 0.003, $OR_{interaction}$= 0.78, 95 % CI = 0.67-0.92 in JHH; $P_{interaction}$ = 0.014, $OR_{interaction}$ = 0.75, 95 % CI = 0.59-0.94 in CGEMS).

Part B: Identify SNP-SNP interaction that without main effects using a novel statistical method of BOOST

After applying the quality control (QC) criteria, there were 509,916 SNPs remained in the analysis in the CGEMS GWAS population. A total of 1,325 pairs of SNP-SNP interactions reached a $P_{interaction}$ cutoff of 1.0E-08. None of the SNP pairs reached a genome-wide significant level of 4.4E-13, if considering 1.25×10^{11} statistical tests. Specifically, 17 pairs of SNP-SNP interaction were found to be significant at a $P_{interaction}$ cutoff of 1.0E-10, and 131 pairs were significant at a $P_{interaction}$ cutoff of 1.0E-09. The most significant hit was observed between rs1178517 located on *CNTN4* gene at chromosome 3 and rs1355096 located on *FAM173B* gene at chromosome 5, with a $P_{interaction}$ of 2.3E-11.

We then examined the interaction effects for the top list of SNP-SNP interactions in another independent GWAS population of JHH. Among the 1,325 pairs of SNP-SNP interactions that were significant at a $P_{interaction}$ of 1.0E-08 in CGEMS, 96 pairs of SNP-SNP interactions were significant at a nominal $P_{interaction}$ of 0.05 in JHH. Sixteen pairs of SNP-SNP interactions were significant at a more stringent $P_{interaction}$ cutoff of 0.01 (**Table 3**). However, no SNP pairs reached a Bonferroni corrected $P_{interaction}$ of 3.8E-05 in JHH population (0.05/1,325).

Among the 16 SNP-SNP interactions that reached the significance level at $P_{interaction}$ < 1.0E-08 for the CGEMS population and $P_{interaction}$ < 0.01 for the JHH population, three were found between SNPs that are both located within the intragenic regions, seven between an intragenic SNP and an intergenic SNP, and six between SNPs both located in intergenic regions. Two interactions deserved to be emphasized because they involve two cancer-related genes. One interaction ($P_{interaction}$ = 5.3E-09 for CGEMS and 1.9E-03 for JHH) was between rs7514217 (within the intron of *PDPN* at 1p36) and rs7934426 (intergenic, at 11p12). The second pair of interaction was between rs11980379, located within the 3'UTR of *IKZF1* at 7p12.2, was found to interact with intergenic rs4314028 at 2p16.3 ($P_{interaction}$ = 5.6E-09 and 3.4E-03, respectively (**Table 3**)

Table 2. Results for top SNPs that interact with the known PCa-risk SNPs ($P_{interaction}$ <1.0E-05 in the meta-analysis, and $P_{interaction}$ <0.05 in each of the three populations)

SNP 1			SNP 2							Meta- results		CAPS		JHH		CGEMS	
CHR	SNP	Gene	CHR	SNP		Minor Allele	Gene	Location	Relative Position	P	OR	P	OR	P	OR	P	OR
2	rs1465618	*THADA*	4	rs11735008	393,303	G	*ABCA11P*	Intergenic	15,921	6.65E-06	0.76(0.68-0.86)	5.04E-03	0.69(0.54-0.90)	3.11E-02	0.83(0.71-0.98)	1.35E-03	0.69(0.55-0.86)
			13	rs9567349	43,535,405	G	*NCRNA00284*	Intergenic	32,806	3.93E-06	0.61(0.49-0.75)	4.57E-04	0.41(0.25-0.67)	3.24E-02	0.72(0.54-0.97)	3.97E-03	0.57(0.39-0.84)
3	rs10934853	*EEFSEC*	9	rs7847271	116,870,633	A	*TNC*	Intron		3.85E-06	0.67(0.56-0.79)	7.73E-03	0.60(0.41-0.87)	2.56E-03	0.68(0.54-0.88)	1.87E-02	0.69(0.50-0.94)
			18	rs998124	40,979,660	G	*MIR4319*	Intergenic	-175,531	5.21E-06	1.33(1.18-1.51)	3.56E-02	1.39(1.02-1.88)	3.64E-03	1.28(1.08-1.51)	3.56E-03	1.42(1.12-1.80)
4	rs17021918	*PDLIM5*	3	rs9757252	86,977,168	T	*VGLL3*	Intergenic	92,645	4.73E-06	1.25(1.13-1.37)	8.45E-03	1.35(1.08-1.69)	2.23E-03	1.22(1.07-1.38)	2.16E-02	1.24(1.03-1.50)
4	rs7679673	*TET2*	9	rs290258	92,595,560	G	*SYK*	Intergenic	-8,273	1.49E-06	0.75(0.67-0.84)	2.11E-03	0.66(0.51-0.86)	3.01E-03	0.78(0.67-0.92)	1.39E-02	0.75(0.59-0.94)
			22	rs5751168	21,175,240	T	*ZNF280B*	Intron		4.11E-06	1.44(1.23-1.67)	4.75E-05	2.19(1.50-3.19)	3.38E-02	1.25(1.02-1.53)	9.09E-03	1.48(1.10-1.99)
7	rs10486567	*JAZF1*	3	rs1795355	41,574,530	T	*ULK4*	Intron		9.46E-06	0.79(0.71-0.88)	3.37E-02	0.77(0.60-0.98)	9.11E-03	0.83(0.72-0.95)	2.14E-03	0.73(0.60-0.89)
			3	rs11720607	174,325,971	G	*SPATA16*	Intron		4.87E-06	0.73(0.63-0.83)	2.45E-03	0.62(0.45-0.84)	2.34E-03	0.75(0.62-0.90)	5.43E-02	0.77(0.58-1.00)
7	rs6465657	*LMTK2*	16	rs8057939	47,951,777	C	*C16orf78*	Intergenic	-13,532	4.71E-06	1.37(1.20-1.57)	3.31E-02	1.43(1.03-1.98)	1.03E-03	1.36(1.13-1.63)	1.70E-02	1.36(1.06-1.75)
8	rs10086908		6	rs10456809	17,921,804	T	*KIF13A*	Intron		4.83E-06	1.25(1.14-1.38)	1.30E-02	1.31(1.06-1.62)	2.28E-03	1.23(1.08-1.40)	1.69E-02	1.26(1.04-1.52)
8	rs1447295	*NA*	7	rs7789197	40,931,652	A	*INHBA*	Intergenic	763,474	3.36E-06	0.66(0.56-0.79)	2.57E-03	0.55(0.38-0.81)	7.27E-03	0.72(0.56-0.91)	1.06E-02	0.66(0.48-0.91)
			9	rs12682851	8,002,418	G	*C9orf123*	Intergenic	-212,619	1.53E-06	0.72(0.63-0.82)	9.66E-03	0.67(0.50-0.91)	2.15E-03	0.75(0.62-0.90)	6.56E-03	0.70(0.54-0.90)
			10	rs10885582	116,317,540	T	*ABLIM1*	Intron		3.70E-06	0.73(0.63-0.83)	9.33E-05	0.54(0.39-0.73)	1.27E-02	0.79(0.66-0.95)	3.46E-02	0.75(0.57-0.98)
8	rs6983267	*NA*	6	rs1011119	19,972,144	G	*ID4*	Intergenic	23,250	7.20E-06	0.81(0.74-0.89)	1.45E-02	0.76(0.61-0.95)	2.85E-03	0.83(0.74-0.94)	1.58E-02	0.80(0.67-0.96)
9	rs1571801	*DAB2IC*	8	rs2219968	79,119,213	A	*PKIA*	Intergenic	-471,678	6.07E-07	1.30(1.17-1.43)	1.18E-02	1.35(1.07-1.70)	1.64E-04	1.30(1.14-1.50)	3.17E-02	1.24(1.02-1.52)
			8	rs13264970	83,236,384	C	*SNX16*	Intergenic	-319,308	3.53E-06	0.77(0.69-0.86)	1.50E-02	0.74(0.59-0.94)	5.02E-03	0.80(0.68-0.93)	3.92E-03	0.72(0.58-0.90)
			10	rs1547851	92,364,806	T	*HTR7*	Intergenic	125,750	7.45E-06	1.59(1.30-1.95)	2.72E-03	1.98(1.27-3.09)	7.35E-03	1.49(1.11-1.99)	2.42E-02	1.52(1.06-2.20)
10	rs4962416	*CTBP2*	5	rs10940579	57,166,575	C	*ACTBL2*	Intergenic	-352,182	3.81E-06	1.32(1.18-1.49)	4.85E-02	1.36(1.00-1.84)	1.77E-03	1.29(1.10-1.51)	4.74E-03	1.39(1.10-1.74)
11	rs10896449	*MYEOV*	12	rs17354197	88,169,501	T	*DUSP6*	Intergenic	96,467	8.82E-06	1.41(1.21-1.64)	3.97E-02	1.49(1.02-2.18)	2.35E-03	1.37(1.12-1.67)	1.10E-02	1.45(1.09-1.92)
11	rs12418451	*MYEOV*	3	rs10513723	176,062,702	A	*NAALADL2*	Intron		5.61E-06	1.41(1.21-1.63)	7.22E-03	1.58(1.13-2.21)	1.61E-02	1.27(1.05-1.54)	1.46E-03	1.64(1.21-2.22)
			8	rs7829048	4,689,690	C	*CSMD1*	Intron		9.76E-06	0.74(0.65-0.85)	1.22E-02	0.68(0.50-0.92)	4.52E-02	0.84(0.71-1.00)	1.81E-04	0.60(0.46-0.78)
			15	rs784411	46,827,089	C	*CEP152*	Intron		1.15E-07	1.42(1.25-1.61)	8.83E-03	1.55(1.12-2.16)	5.28E-04	1.34(1.14-1.58)	1.32E-03	1.53(1.18-1.99)
11	rs7127900	IGF2, IGF2AS	8	rs13258681	124,783,903	C	*ANXA13*	Intron		3.65E-06	1.32(1.17-1.48)	4.58E-02	1.32(1.01-1.73)	5.43E-04	1.33(1.13-1.56)	1.90E-02	1.31(1.05-1.64)
		INS,TH	22	rs12628051	38,984,222	C	*TNRC6B*	Intron		3.39E-06	1.30(1.17-1.46)	1.82E-03	1.50(1.16-1.93)	6.14E-03	1.24(1.06-1.44)	1.44E-02	1.32(1.06-1.65)
17	rs4430796	*HNF1B*	1	rs731174	37,969,428	C	*EPHA10*	Intron		4.55E-06	1.27(1.15-1.40)	5.03E-02	1.24(1.00-1.54)	3.20E-02	1.18(1.01-1.38)	1.13E-04	1.41(1.19-1.68)
			2	rs12694942	158,518,681	T	*UPP2*	Intergenic	-41,256	7.84E-06	0.80(0.73-0.88)	5.22E-03	0.75(0.61-0.92)	1.87E-03	0.79(0.68-0.92)	6.36E-02	0.85(0.72-1.01)
			9	rs10812303	25,712,117	T	*TUSC1*	Intergenic	-43,261	5.59E-06	1.34(1.18-1.52)	6.63E-03	1.46(1.11-1.91)	7.57E-03	1.27(1.07-1.52)	9.01E-03	1.37(1.08-1.73)
17	rs1859962	*NA*	2	rs16867225	180,749,531	A	*CWC22*	Intergenic	-169,506	3.12E-06	0.64(0.53-0.77)	1.33E-03	0.48(0.30-0.75)	4.80E-03	0.70(0.54-0.90)	1.84E-02	0.65(0.45-0.93)

			7	rs10277209	108,790,810	C	*C7orf66*	Intergenic	478,937	3.81E-06	1.36(1.19-1.55)	4.08E-02	1.39(1.01-1.89)	4.04E-03	1.29(1.08-1.53)	1.55E-03	1.52(1.17-1.97)
19	rs2735839	*KLK3*	20	rs6089829	61,139,481	A	*LOC63930*	Intron		3.21E-06	0.74(0.65-0.84)	2.53E-02	0.71(0.52-0.96)	1.26E-03	0.75(0.64-0.90)	1.10E-02	0.72(0.56-0.93)
19	rs8102476	*PPP1R14A*	1	rs1866967	29,958,249	G	*PTPRU*	Intergenic	432,337	5.16E-06	0.82(0.75-0.89)	2.70E-02	0.79(0.64-0.97)	5.10E-03	0.85(0.75-0.95)	2.97E-03	0.78(0.66-0.92)
19	rs887391	*NA*	4	rs735172	5,809,770	C	*EVC*	Intron		2.03E-06	1.31(1.17-1.46)	6.58E-03	1.43(1.10-1.85)	2.37E-03	1.27(1.09-1.47)	1.02E-02	1.32(1.07-1.63)
			5	rs4463179	13,558,432	A	*DNAH5*	Intergenic	185,005	2.22E-06	0.64(0.53-0.77)	3.00E-02	0.64(0.42-0.96)	8.35E-05	0.59(0.46-0.77)	8.28E-02	0.73(0.51-1.04)
22	rs9623117	*TNRC6B*	4	rs1713511	43,472,127	A	*KCTD8*	Intergenic	398,550	7.87E-06	1.31(1.16-1.47)	4.94E-03	1.54(1.14-2.09)	6.71E-03	1.23(1.06-1.44)	1.06E-02	1.36(1.07-1.72)

Abbreviations: SNP1 indicate the 32 known PCa-risk SNPs. SNP 2 indicates the interacting SNPs; Chr, chromosome; Relative position is the distance of SNP2 relative to the nearest gene if SNP2 is located in the intergenic region; OR : Odds Ratio; P and OR are for the multiplicative interaction term. P and OR for the Meta-analysis are calculated based on a Cochran-Mantel-Haenszel test. CAPS: PCa case-control study from Sweden; JHH: Johns Hopkins Hospital; CGEMS: the Cancer Genetic Markers of Susceptibility;

Table 3. Results for top list of SNP-SNP interactions in the CGEMS and JHH populations ($P_{interaction}$ <1.0E-08 in CGEMS and $P_{interaction}$ <0.01 in JHH).

	SNP A						SNP B						Interaction	Annotation for SNP A			Annotation for SNP B		
Study	rsID	CHR	BP	A1	MAF	Main effect P	rsID	CHR	BP	A1	MAF	Main effect P	BOOST P	Location	Gene	Relative Position	Location	Gene	Relative Position
CGEMS	rs7514217	1	13,795,386	G	0.42	0.75	rs7934426	11	37,270,065	G	0.46	0.38	5.3E-09	Intron	*PDPN*	.	Intergenic	*RAG2*	-693,660
JHH				G	0.45	0.16				G	0.44	0.37	1.9E-03						
CGEMS	rs2503220	1	66,272,145	G	0.08	0.80	rs2579790	10	77,725,730	C	0.38	0.32	9.3E-09	Intron	*PDE4B*	.	Intron	*C10orf11*	.
JHH				G	0.08	0.75				C	0.39	0.23	3.1E-03						
CGEMS	rs13402702	2	29,033,906	G	0.2	0.05	rs7329899	13	105,276,784	A	0.26	0.99	6.4E-09	Intergenic	*SNORD53*	30,394	Intergenic	*DAOA*	335,400
JHH				G	0.18	0.15				A	0.25	0.33	9.0E-03						
CGEMS	rs4314028	2	51,990,169	C	0.25	0.41	rs11980379	7	50,437,475	C	0.26	0.49	5.6E-09	Intergenic	*CHAC2*	-1,858,264	3' UTR	*IKZF1*	.
JHH				C	0.26	0.99				C	0.26	0.18	3.4E-03						
CGEMS	rs4314028	2	51,990,169	C	0.25	0.41	rs4132601	7	50,438,098	G	0.26	0.51	3.4E-09	Intergenic	*CHAC2*	-1,858,264	3' UTR	*IKZF1*	.
JHH				C	0.26	0.99				G	0.26	0.14	5.1E-03						
CGEMS	rs4664789	2	156,575,782	C	0.49	0.97	rs8019172	14	51,715,669	A	0.1	0.48	8.8E-09	Intergenic	*NR4A2*	313,408	Intergenic	*PTGDR*	-88,512
JHH				A	0.49	0.17				A	0.1	0.35	5.8E-03						
CGEMS	rs4973194	2	229,945,721	G	0.37	0.49	rs1949403	3	6,070,643	C	0.28	0.26	9.4E-09	Intron	*DNER*	.	Intergenic	*EDEM1*	833,993
JHH				G	0.36	0.91				C	0.27	0.65	8.9E-03						
CGEMS	rs6772801	3	140,979,161	G	0.29	0.16	rs6955437	7	137,547,448	T	0.11	0.77	3.9E-09	Intergenic	*NMNAT3*	-99,631	Intergenic	*AKR1D1*	93,858
JHH				G	0.27	3.3E-04				T	0.09	0.31	8.9E-03						
CGEMS	rs6878100	5	129,370,204	A	0.33	0.77	rs2960753	7	141,386,546	T	0.4	0.47	3.6E-09	Intron	*CHSY3*	.	Intron	*MGAM*	.
JHH				A	0.33	0.44				T	0.39	0.39	9.7E-03						
CGEMS	rs6948622	7	145,043,041	A	0.36	0.85	rs1154140	14	40,456,756	G	0.3	0.23	8.2E-09	Intergenic	*TPK1*	-878,962	Intergenic	*LRFN5*	-689,758
JHH				A	0.34	0.28				G	0.32	0.14	9.2E-03						
CGEMS	rs12682543	8	29,135,477	G	0.34	0.04	rs11231168	11	62,158,196	T	0.42	0.23	2.2E-09	Intron	*KIF13B*	.	Intron	*GANAB*	.
JHH				G	0.32	0.12				T	0.4	0.34	4.7E-03						
CGEMS	rs10810961	9	18,361,966	G	0.10	0.11	rs643853	21	43,656,244	A	0.18	0.05	3.7E-09	Intergenic	*MIR3152*	-201,338	Intergenic	*SIK1*	2,583
JHH				G	0.11	0.26				A	0.21	0.22	5.2E-03						
CGEMS	rs4837960	9	124,163,792	T	0.15	0.98	rs275769	12	123,687,425	T	0.31	0.45	9.8E-09	Intergenic	*PTGS1*	-9,258	Intergenic	*SCARB1*	140,702
JHH				T	0.14	0.45				T	0.3	0.24	8.9E-03						
CGEMS	rs1038972	10	31,243,098	T	0.11	0.17	rs2022896	14	27,074,470	A	0.24	0.65	4.2E-09	Intron	*ZNF438*	.	Intergenic	*MIR4307*	626,699
JHH				T	0.12	0.33				A	0.25	0.70	7.8E-03						
CGEMS	rs12861843	13	35,638,422	C	0.42	0.06	rs3862743	13	41,208,633	C	0.39	0.34	4.1E-09	Intergenic	*SOHLH2*	1,923	Intron	*KIAA0564*	.
JHH				C	0.45	0.91				C	0.4	0.28	3.2E-03						
CGEMS	rs2136267	13	107,332,783	T	0.25	0.37	rs1884393	20	1,404,079	A	0.11	0.46	9.3E-09	Intergenic	*LIG4*	325,010	3' UTR	*SIRPB2*	.
JHH				T	0.27	0.33				A	0.11	0.95	6.5E-03						

Abbreviations: SNP A indicates the first interacting SNP; SNP B indicates the second interacting SNP; .

CHR, chromosome; MAF, minor allele frequency;

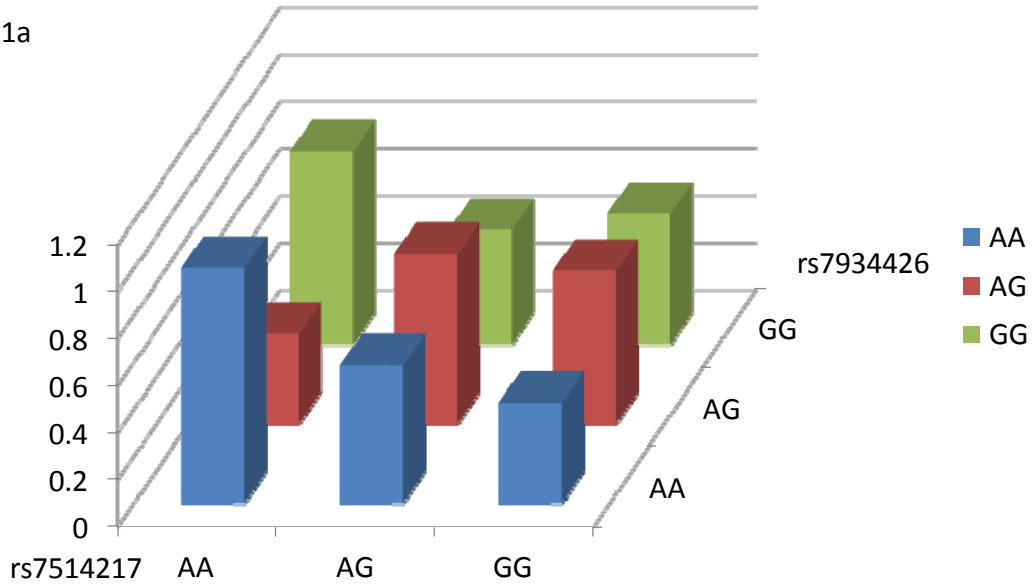
Main effect P refers to the single-locus P-value based on the two-degree of freedom test; BOOST P refers to the P-values for the multiplicative interaction term, as calculated by BOOST approach;

Relative position is the distance of SNPA/SNPB relative to the nearest gene if SNPA/SNPB is located in the intergenic region;

CGEMS: the Cancer Genetic Markers of Susceptibility; JHH: Johns Hopkins Hospital;

We then carefully examined the two-locus interaction pattern of the above two SNP pairs in CGEMS and JHH populations. Figure 1a showed the odds ratios for the 9 combinations of the genotypes of rs7514217 and rs7934426. Men who carried “GG/GG” double homozygotes for both SNPs had a significantly decreased risk of developing PCa (OR = 0.56, 95% CI= 0.34-0.91; *P* = 0.02) in CGEMS compared with men who carried homozygous “A” allele (major allele) for both rs7514217 and rs7934436 (reference group) (Figure 1a). A similar pattern of interaction was observed in JHH population. Particularly, men who carry GG/GG genotypes also had a decreased risk of developing PCa (OR = 0.51, 95% CI = 0.36-0.74; *P* = 4.0E-04, Figure 1b) compared with men who carried homozygous “A” allele (major allele) for both of the SNPs. Figure 2a and 2b showed the interaction pattern between rs11980379 and rs4314028 in CGEMS and JHH, respectively. Men who carried the “TC/CC” genotype for rs11980379 and rs4314028 had a marginal significantly increased risk for PCa (OR = 1.86, 95% CI = 0.98-3.64; *P* = 0.05) in CGEMS, compared with men with homozygous “T” allele (major allele) (Figure 2a). Similar in JHH, men who carry the “TC/CC” genotype combination also had a significantly increased risk for PCa (OR = 1.83, 95% CI = 1.24-2.72; *P* = 2.1E-03) (Figure 2b). However, the interaction pattern was not consistent with a dominant model since “TC/CC” genotype for rs11980379 and rs4314028 showed an increase in risk while the “CC/CC” genotype shows a reduced risk (both compared to “ TT/TT”) (Figure 2b).

Figure 1a

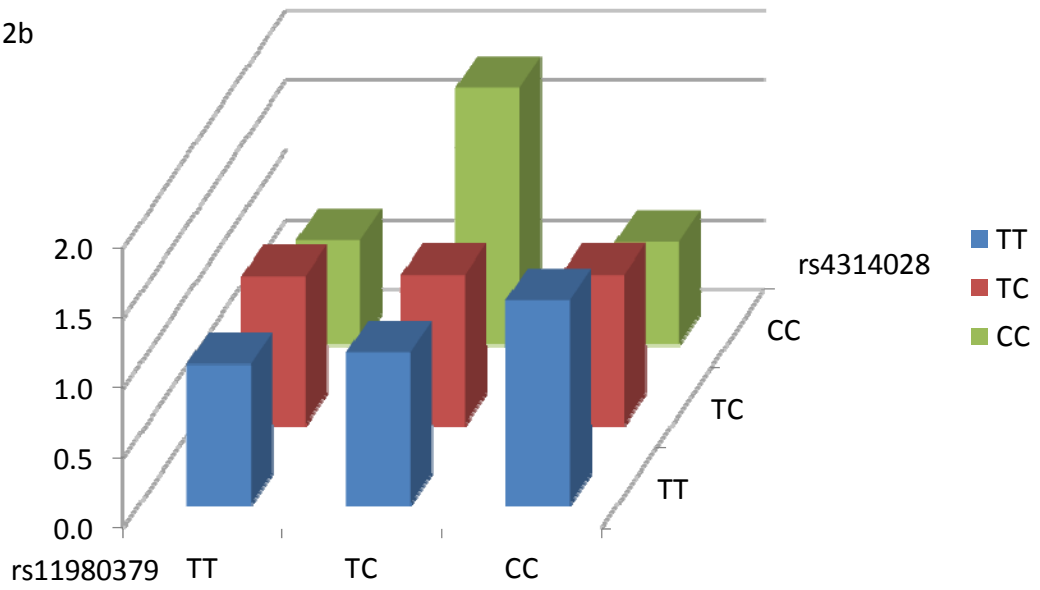


Odds ratio (95% confidence interval), and P-values for SNP pair of rs7514217 (column) and rs7934426 (row)

	AA	AG	GG
AA	1.000	0.38 (0.27, 0.54) 1.7E-08	0.82 (0.55, 1.23) 0.36
AG	0.59 (0.42, 0.83) 2.5E-3	0.72 (0.53, 0.99) 0.05	0.49 (0.33, 0.71) 2E-04
GG	0.43 (0.27, 0.67) 3.0E-4	0.65 (0.44, 0.95) 0.03	0.56 (0.34, 0.91) 0.02

Figure 1a and 1b. Interaction between rs7514217 and rs7934426 in CGEMS (1a) and JHH (1b). X-axis represent the genotype for rs7514217, Y-axis represent the genotype for rs7934426, the Z-axis showed the odds ratio of the SNP pair (rs7514217 and rs7934426). Odds ratios are estimated relative to the baseline AA/AA double homozygote.

Figure 2b



Odds ratio (95% confidence interval), and P-values for SNP pair of rs4314028 (column) and rs11980379 (row)

	TT	TC	CC
TT	**1.00**	**1.06 (0.90,1.25) 0.50**	**0.75 (0.52, 1.06) 0.10**
TC	**1.08 (0.92, 1.28) 0.34**	**1.07 (0.89, 1.29) 0.48**	**1.83 (1.24, 2.72) 2.1E-03**
CC	**1.46 (1.06, 2.00) 0.02**	**1.07 (0.75, 1.51) 0.73**	**0.73 (0.29, 1.69) 0.57**

Figure 2a and 2b. Interaction between rs11980379 and rs4314028 in CGEMS (2a) and JHH (2b). X-axis represent the genotype for rs11980379, Y-axis represent the genotype for rs4314028, the Z-axis showed the odds ratio of the SNP pair (rs11980379 and rs4314028). Odds ratios are estimated relative to the baseline AA/AA double homozygote.

Discussion

Part A: Identify SNPs that interact with 32 known PCa risk variants.

To our knowledge, our study represents one of the first comprehensive gene-gene interaction scans in three PCa GWAS populations. Specifically, we performed a genome-wide gene-gene interaction scan for each of the 32 known prostate cancer risk-associated variants identified from genome-wide association studies in three case-control populations of European descents, which includes a total of 4,723 PCa cases and 4,792 controls. In the meta-analysis, we found 35 pairs of SNP-SNP interactions that were significantly associated with PCa risk ($P_{interaction}$ <1E-05). In addition, the interactions for those 35 pairs were significant in all three populations (all $P_{interaction}$ <0.05). Among those 35 pairs of statistically significant interactions, we emphasized three pairs of interactions with potential biological implication, including an interaction between rs12418451 in *MYEOV* and rs16961635 in *CEP152*, with a $P_{interaction}$ of 1.15E-07 (OR = 1.42, 95 % CI = 1.25-1.61), an interaction between rs7127900 at *IGF2/IGF2AS* region and rs12628051 in the intron of *TNRC6B*, with a $P_{interaction}$ of 3.39E-06 (OR = 1.30, 95% CI = 1.17-1.46), an interaction between rs7679763 in *TET2* gene region and rs290258 in the promoter region of SYK, with a $P_{interaction}$ of 1.49E-06 (OR = 0.75, 95% CI = 0.67-0.84).

The discovery of approximately three dozen PCa risk variants using single SNP analysis suggests that it is possible to detect individual risk variants. However, when the underlying genetic model involves interaction of multiple genes, a single gene approach is less effective and may not be able to explain the complex etiology of the disease. Therefore, evaluation of the joint effect (epistasis) of multiple genetic variants is critical to understand the underlying causes of complex diseases [22], especially in the situation where several individual risk variants have been identified. The next question is to explore whether other SNPs interact with those SNPs to modify risk to PCa. The identified loci that interact with the known PCa risk-associated SNPs may help to elucidate the underlying molecular mechanisms of the associations of those risk SNPs.

The most significant interaction was seen between the PCa risk-associated SNPS rs12418451 and rs784411. The SNP rs12418451 is located at the 11q13.2 that is ~77kb upstream of *TPCN2*, a putative cation-selective ion channel gene, and ~126kb upstream of *MYEOV*, an oncogene that has been implicated in multiple cancers [23-27]. The SNP rs784411 resides in the intron of CEP152, a centrosomal protein that was recently shown to function as a regulator of genomic integrity [28] and cellular response to DNA damage[29]. Given the limited information, we speculate that observed interaction may reflect the close collaboration of MYEOV (or TPCN2, even though it is less likely) and CEP152 in the same or different oncogenic pathways that drive the tumorigenesis of prostatic epithelial cells.

Among the two SNPs that were shown to consistently interact with the PCa risk-associated SNP rs7127900 at 11p15.5, one SNP (rs12628051) is located within TNRC6B, which encodes a RNA interference (RNAi) machinery component protein crucial for the miRNA/siRNA-dependent translational repression or degradation of target mRNAs. It is worthy to mention that this gene also contains a GWAS-identified PCa risk-associated SNP (rs9623117). Several mechanisms may potentially explain for these interactions. Firstly, we noticed that at ~70 kb telomeric to rs7127900 reside the PCa-implicated *IGF2* gene and its antisense transcript-encoding *IGF2AS*. IGF2 encodes a member of the insulin family of polypeptide growth factors that promotes cell proliferation during fetal development but becomes less active in healthy adults due to genomic imprinting. Dysregulated overexpression of IGF2 caused by loss of imprinting (LOI) has been associated with a variety of human cancers including PCa [30-33]. IGF2AS encodes a predictably non-coding RNA that is antisense to IGF2 and thus may potentially regulate IGF2 expression through RNAi in a similar manner as some other natural antisense transcripts. Thus one plausible scenario is that TNRC6B may affect the RNAi-mediated transcriptional regulation of IGF2AS on IGF2, which may underlie the observed interaction between genetic variants within these two loci. Secondly, there are two microRNA (miRNA) genes located at 11p15.5, miR-4686 (~40kb from the PCa-risk SNP rs7127900) and miR-483 (~80kb from rs7127900). Although the role of miR-4686 remains to be determined, miR-483 has been demonstrated to act as an oncogene to suppress proapoptotic BBC3 (PUMA) or tumor suppressive DPC4(Smad4) in a variety of human cancers[34,35]. Thus an alternative mechanism for the observed interaction between the 11p15.5 locus and the TNRC6B locus is that genetic variants in TNRC6B may affect the miR-483 (or miR4686)-mediated RNAi toward its/their target tumor suppressor genes.

Another pair of interacting SNPs were found between rs7679673 (~ 6kb upstream of *TET2*) and rs290258 (~8kb upstream of *SYK*). TET2 encodes an enzyme hydroxylating methylcytosine and is implicated in epigenetic programming that involves DNA methylation and demethylation (Reviewed in [36]). The critical role of TET2 in cancer is suggested by the observation that loss of function mutations of TET2 are frequently identified in various hematologic malignancies[37,38]. As a non-receptor Tyrosine protein kinase that mediates cellular proliferation and differentiation, SYK is believed to function as a potential tumor suppressive gene (reviewed in [39]. It is noteworthy that

hypermethylation of SYK gene promoter has been frequently found in and widely associated with lung, gastric, and breast cancer [40,41]. Thus although it remains to be determined whether SYK promoter in prostatic tumors also undergoes silencing via DNA methylation, the observed interaction between TET2 and SYK suggests that it is a plausible hypothesis.

Two SNPs (rs731174 and rs10812303) were found to interact with the GWAS-identified PCa risk-associated SNP rs4430796, residing within *HNF1B*, a homeodomain-containing transcription factor whose expression alteration has been widely implicated in various human cancers including PCa. The SNP rs731174 is located within the intron of *EPHA10*, a member of the EPH subfamily of receptor tyrosine kinases (RTKs). This family of RTKs play an important role in cell-cell communication regulating cell attachment, shape, and mobility in epithelial cells and are believed to be implicated in carcinogenesis (reviewed in [42]. It is possible that HNF1Ba and EPHA10 collaborate in the signaling network that is crucial for the well-being of prostatic cells whereas the genetic variants located within these two genes may synergistically contribute to the oncogenesis of PCa. The other SNP rs10812303 is ~40kb upstream of *TUSC1*, an intronless gene that has been suggested to serve as a tumor suppressor in lung tumorigenesis [43].Thus the interaction between genetic variants in TUSC1 and HNF1B may also suggest a plausible collaboration of these two genes.

In summary, our systematic evaluation of gene-gene interactions in three GWAS populations suggested a list of loci interacting with known PCa risk-associated SNPs that may warrant follow-up in other study populations. Three pairs of interactions are worthwhile to be emphasized, including an interaction between rs12418451 in the *MYEOV* gene region and rs784411 in the intron of *CEP152*, an interaction between rs7127900 in the *IGF2/IGF2AS* gene region and rs12628051 in the intron of *TNRC6B*, and an interaction between rs7679673 in the *TET2* gene region and rs290258 in the intron of *SYK.* Those results showed statistical evidence for genes interacting with known risk-associated SNPs on PCa risk. The interacting loci identified also provide more hints on the underlying molecular mechanism of the associations with PCa risk for the known risk-associated SNPs.

Part B: Identify SNP-SNP interaction that without main effects using a novel statistical method of BOOST

The SNP-SNP interactions revealed by our study involved loci with no evidence of main effect or weak marginal effect. Previous studies also showed that all the SNPs that were implicated in the top hits of interactions in Type 1 Diabetes and Type 2 Diabetes displayed weak main effects [44,45]. Culverhouse et al. also reported large interaction effects in the complete absence of marginal effect [46]. Thus these findings, combined with ours, highlight the need for exhaustive search when evaluating epistasis in a genome-wide scale. To reduce the computation burden, many methods have been developed with pre-screening algorithms built in [22,47,48]. However, most of the pre-screening algorithms are based on evaluating the marginal effects of single-locus. The SNPs with weak or no marginal effects but with significant interaction effects will be screened out based on such algorithms. In contrast, BOOST was able to evaluate all pair-wise interactions on the genome-wide scale with a relatively fast speed, which allows for exhaustive search of interaction effects across the entire genome.

Specifically, we want to emphasize on two pairs of interactions implicated in our study because they involve two cancer-related genes. One interaction was between rs7514217 (within the intron of *PDPN* at 1p36) and rs7934426 (intergenic, at 11p12). *PDPN* encodes a mucin-type transmembrane glycoprotein (podoplanin) that reportedly plays diverse functions including regulating actin

cytoskeleton organization and cell migration[49,50], inducing platelet aggregation[51], and modulating lymphatic vasculature formation[52]. The crucial role of *PDPN* in tumorigenesis is indicated by the observations that *PDPN* is frequently overexpressed in various human cancer types[53-56], and that podoplanin facilitates tumor cell migration and invasion [57] and promotes lymphangiogenesis and lymph node metastasis[58]. Expression of PDPN was also reported in prostatic tumors [59], suggesting a role in PCa as well.

The second pair of interaction was between rs11980379, located within the 3'UTR of *IKZF1* at 7p12.2, and rs4314028, located on an intergenic region at 2p16.3. *IKZF1* encodes a Ikaros family zinc finger transcription factor that normally directs hematopoietic lineage commitment and pituitary neuroendocrine cell expansion by regulating differentiation, proliferation, and apoptosis of these cell lineages[60,61]. *IKZF1* is considered a hematological and pituitary tumor suppressor such that abnormalities in its splicing have been associated with leukemias[62] and pituitary tumors[63]. Recently expression of *IKZF1* was also found in a variety of other human tissues including prostate and was associated with prognosis of breast, lung, ovarian and skin cancers [64], suggesting a potentially important role of *IKZF1* in other cancer types such as PCa. However, given the "gene desert" localizations of the partner SNPs that interact with these *PDPN/ IKZF1*-harbored SNPs, it is hard to make biological inferences as what molecular mechanisms might account for these two interactions. Nonetheless, as revealed by the ongoing ENCODE project and several genome-wide cistrome studies on important transcription factors such as Forkhead box A1 (FOXA1)[65], Androgen Receptor (AR) [65] and Estrogen Receptor (ER) [66], the DNA sequences whereby certain functionally critical transcription factors bind and regulate expression of their target genes are extensively localized throughout the genome, which sometimes are several hundred kbps away from their target genes and may reside in intergenic regions (enhancers or suppressors). Thus it is possible that these intergenic SNPs, or those in Linkage Disequilibrium (LD) with them, may lie within the DNA sequences containing enhancer or suppressor activities that distantly regulate one or several target genes. These genes may collaborate with *PDPN/IKZF1* in a common signaling network that combinatorially determines the well-being of prostatic epithelial cells. A certain combination of genetic variants in these interacting intergenic and intragenic loci, though insufficient by themselves alone, may cause the synthetic deficiency of the crucial signaling network and result in explicitly increased risk for PCa. This enhancer/suppressor-involved mechanism may potentially provide explanations for the remaining interactions as well, especially for the six interactions which involve two intergenic SNPs. However, it should be pointed out these contemplations are largely speculative and may require in-depth mechanistic and functional studies to prove.

In summary, we systematically evaluated genome-wide SNP-SNP interactions using a novel statistical approach named BOOST. We identified 1,325 pairs of SNP-SNP interactions with a $P_{\text{interaction}}$ cutoff of 1.0E-08 in the discover population of CGEMS. Among the 1,325 pairs of interactions, 16 pairs of interactions were also significant in an independent population from JHH, at a $P_{\text{interaction}}$ cutoff of 0.01. Our study represents one of the first application studies which implemented a novel statistical method of BOOST to detect interactions on a genome-wide scale. The pairs of SNP-SNP interactions suggested in our study represent the first step towards obtaining further biological insight into the high-dimensional etiology of prostate cancer.

KEY RESEARCH ACCOMPLISHMENTS

1) Completed a genome-wide search for SNPs that interact with the 32 known risk SNPs in the CAPS population

2) Completed a genome-wide search for SNPs that interact with the 32 known risk SNPs in the JHH population
3) Completed a genome-wide search for SNPs that interact with the 32 known risk SNPs in the CGEMS population
4) Completed a meta-analysis and a fine-mapping study of the three GWAS populations (comprised of 4,723 PCa cases and 4,792 controls) to identify SNPs that interact with the 32 risk SNPs
5) Completed an exhaustive search of genome-wide SNP-SNP interactions using a novel statistical approach of BOOST in the JHH and CGEMS populations
6) Published two manuscripts summarizing the key research findings of the funded project in high-quality peer-reviewed journals

REPORTABLE OUTCOMES

A. SNPs that interacted with known PCa risk loci
   1) Thirty-five pairs of SNP-SNP interactions were significantly associated with PCa risk (a $P_{interaction}$ < 1E-05) in the meta-analysis. In addition, the interaction for those 35 pairs was significant in all three populations (all $P_{interaction}$ < 0.05 in CGEMS, JHH, and CAPS) (see Table 2).
   2) The most significant interaction was detected between rs12418451 in *MYEOV* and rs784411 in *CEP152*, with a $P_{interaction}$ of 1.15E-07 in the meta-analysis of three populations (Table 2).
   3) Two additional pairs of interactions that were significant at a $P_{interaction}$ < of 1E-05 in the meta-analysis were biologically interesting, including an interaction between rs7127900 at *IGF2/IGF2AS* region and rs12628051 in the intron of *TNRC6B*, with a $P_{interaction}$ of 3.39E-06 (OR = 1.30, 95% CI = 1.17-1.46), and an interaction between rs7679763 in *TET2* gene region and rs290258 in the promoter region of SYK, with a $P_{interaction}$ of 1.49E-06 (OR = 0.75, 95% CI = 0.67-0.84).

B. SNP-SNP interactions without main effects
   1) Among the 1,325 pairs of SNP-SNP interactions with a $P_{interaction}$ cutoff of 1.0E-08 in the discovery population of CGEMS, 16 pairs of interactions were also significant in an independent population from JHH, at a $P_{interaction}$ cutoff of 0.01. (Table 3)
   2) Two interactions deserve to be emphasized because they involve two cancer-related genes. One interaction ($P_{interaction}$ = 5.3E-09 for CGEMS and 1.9E-03 for JHH) was between rs7514217 (within the intron of *PDPN* at 1p36) and rs7934426 (intergenic, at 11p12). The second pair of interaction was between rs11980379, located within the 3'UTR of *IKZF1* at 7p12.2, was found to interact with intergenic rs4314028 at 2p16.3 ($P_{interaction}$ = 5.6E-09 and 3.4E-03, respectively. (Figure 1 and Figure 2)

CONCLUSION

1) We have achieved all the goals described in the approved statement of work.
2) We have identified and confirmed SNPs in the genome that significantly interact with the 32 known PCa risk SNPs in three study populations.

3) The interacting loci identified provide more hints into the underlying molecular mechanism of the associations with PCa risk for the known risk-associated SNPs.
4) We have also identified and confirmed SNP-SNP interactions without main effects which confer increased risk for prostate cancer.
5) Our study represents the first step towards obtaining further biological insight into the high-dimensional etiology of prostate cancer.

Published papers in the funded period (Supported by this grant)

1. Tao S, Wang Z, Feng J, Hsu FC, Jin G, Kin ST, Zhang Z, Gronberg H, Zheng, SL, Isaacs WB, XU J, **Sun J**. A Genome-Wide Search for Loci Interacting with Known Prostate Cancer Risk-Associated Genetic Variants. Carcinogenesis . 2012, 33(3):598-603.

2. Tao S, Feng J, Webster T, Jin G, Hsu FC, Chen SH, Zhang Z, Zheng, SL, Isaacs WB, Xu J, **Sun J**. Genome-wide Two-Locus Epistasis Scan in Prostate Cancer Using Two European Populations. Hum Genet. 2012 , 131(7):1225-34.

References

1. **Xu, J., Zheng, S.L., Isaacs, S.D., Wiley, K.E., Wiklund, F., Sun, J., Kader, A.K., Li, G., Purcell, L.D., Kim, S.T., Hsu, F.C., Stattin, P., Hugosson, J., Adolfsson, J., Walsh, P.C., Trent, J.M., Duggan, D., Carpten, J., Gronberg, H. and Isaacs, W.B. (2010) Inherited genetic variant predisposes to aggressive but not indolent prostate cancer. *Proc Natl Acad Sci U S A*, 107, 2136-40.**
2. **Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*, 39, 906-13.**
3. **Amundadottir, L.T., Sulem, P., Gudmundsson, J., Helgason, A., Baker, A., Agnarsson, B.A., Sigurdsson, A., Benediktsdottir, K.R., Cazier, J.B., Sainz, J., Jakobsdottir, M., Kostic, J., Magnusdottir, D.N., Ghosh, S., Agnarsson, K., Birgisdottir, B., Le Roux, L., Olafsdottir, A., Blondal, T., Andresdottir, M., Gretarsdottir, O.S., Bergthorsson, J.T., Gudbjartsson, D., Gylfason, A., Thorleifsson, G., Manolescu, A., Kristjansson, K., Geirsson, G., Isaksson, H., Douglas, J., Johansson, J.E., Balter, K., Wiklund, F., Montie, J.E., Yu, X., Suarez, B.K., Ober, C., Cooney, K.A., Gronberg, H., Catalona, W.J., Einarsson, G.V., Barkardottir, R.B., Gulcher, J.R., Kong, A., Thorsteinsdottir, U. and Stefansson, K. (2006) A common variant associated with prostate cancer in European and African populations. *Nat Genet*, 38, 652-8.**
4. **Gudmundsson, J., Sulem, P., Manolescu, A., Amundadottir, L.T., Gudbjartsson, D., Helgason, A., Rafnar, T., Bergthorsson, J.T., Agnarsson, B.A., Baker, A., Sigurdsson, A., Benediktsdottir, K.R., Jakobsdottir, M., Xu, J., Blondal, T., Kostic, J., Sun, J., Ghosh, S., Stacey, S.N., Mouy, M., Saemundsdottir, J., Backman, V.M., Kristjansson, K., Tres, A., Partin, A.W., Albers-Akkers, M.T., Godino-Ivan Marcos, J., Walsh, P.C., Swinkels, D.W., Navarrete, S., Isaacs, S.D., Aben, K.K., Graif, T., Cashy, J., Ruiz-Echarri, M., Wiley, K.E., Suarez, B.K., Witjes, J.A., Frigge, M., Ober, C., Jonsson, E., Einarsson, G.V., Mayordomo, J.I., Kiemeney, L.A., Isaacs, W.B., Catalona, W.J., Barkardottir, R.B., Gulcher, J.R., Thorsteinsdottir, U., Kong, A. and Stefansson, K. (2007) Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat Genet*, 39, 631-7.**
5. **Yeager, M., Orr, N., Hayes, R.B., Jacobs, K.B., Kraft, P., Wacholder, S., Minichiello, M.J., Fearnhead, P., Yu, K., Chatterjee, N., Wang, Z., Welch, R., Staats, B.J., Calle, E.E., Feigelson, H.S., Thun, M.J., Rodriguez, C., Albanes, D., Virtamo, J., Weinstein, S., Schumacher, F.R., Giovannucci, E., Willett, W.C., Cancel-Tassin, G., Cussenot, O., Valeri, A., Andriole, G.L., Gelmann, E.P., Tucker, M., Gerhard, D.S., Fraumeni, J.F., Jr., Hoover, R., Hunter, D.J., Chanock, S.J. and Thomas, G. (2007) Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet*, 39, 645-9.**
6. **Gudmundsson, J., Sulem, P., Steinthorsdottir, V., Bergthorsson, J.T., Thorleifsson, G., Manolescu, A., Rafnar, T., Gudbjartsson, D., Agnarsson, B.A., Baker, A., Sigurdsson, A., Benediktsdottir, K.R., Jakobsdottir, M., Blondal, T., Stacey, S.N., Helgason, A., Gunnarsdottir, S., Olafsdottir, A., Kristinsson, K.T., Birgisdottir, B., Ghosh, S., Thorlacius, S., Magnusdottir, D., Stefansdottir, G., Kristjansson, K., Bagger, Y., Wilensky, R.L., Reilly, M.P., Morris, A.D., Kimber, C.H., Adeyemo, A., Chen, Y., Zhou, J., So, W.Y., Tong, P.C., Ng, M.C., Hansen, T., Andersen, G., Borch-Johnsen, K., Jorgensen, T., Tres, A., Fuertes, F., Ruiz-Echarri, M., Asin, L., Saez, B., van Boven, E., Klaver, S., Swinkels, D.W., Aben, K.K., Graif, T., Cashy, J., Suarez, B.K., van Vierssen Trip, O., Frigge, M.L., Ober, C., Hofker, M.H., Wijmenga, C., Christiansen, C., Rader, D.J., Palmer, C.N., Rotimi, C., Chan, J.C., Pedersen, O., Sigurdsson, G., Benediktsson, R., Jonsson, E., Einarsson, G.V., Mayordomo, J.I., Catalona, W.J., Kiemeney, L.A., Barkardottir, R.B., Gulcher, J.R., Thorsteinsdottir, U., Kong, A. and Stefansson, K. (2007) Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat Genet*, 39, 977-83.**
7. **Duggan, D., Zheng, S.L., Knowlton, M., Benitez, D., Dimitrov, L., Wiklund, F., Robbins, C., Isaacs, S.D., Cheng, Y., Li, G., Sun, J., Chang, B.L., Marovich, L., Wiley, K.E., Balter, K., Stattin, P., Adami, H.O., Gielzak, M., Yan, G., Sauvageot, J., Liu, W., Kim, J.W., Bleecker, E.R., Meyers, D.A., Trock, B.J., Partin, A.W., Walsh, P.C., Isaacs, W.B., Gronberg, H., Xu, J. and Carpten, J.D.**

(2007) Two genome-wide association studies of aggressive prostate cancer implicate putative prostate tumor suppressor gene DAB2IP. *J Natl Cancer Inst,* 99, 1836-44.

8. **Thomas, G., Jacobs, K.B., Yeager, M., Kraft, P., Wacholder, S., Orr, N., Yu, K., Chatterjee, N., Welch, R., Hutchinson, A., Crenshaw, A., Cancel-Tassin, G., Staats, B.J., Wang, Z., Gonzalez-Bosquet, J., Fang, J., Deng, X., Berndt, S.I., Calle, E.E., Feigelson, H.S., Thun, M.J., Rodriguez, C., Albanes, D., Virtamo, J., Weinstein, S., Schumacher, F.R., Giovannucci, E., Willett, W.C., Cussenot, O., Valeri, A., Andriole, G.L., Crawford, E.D., Tucker, M., Gerhard, D.S., Fraumeni, J.F., Jr., Hoover, R., Hayes, R.B., Hunter, D.J. and Chanock, S.J. (2008) Multiple loci identified in a genome-wide association study of prostate cancer. *Nat Genet,* 40, 310-5.**
9. **Gudmundsson, J., Sulem, P., Rafnar, T., Bergthorsson, J.T., Manolescu, A., Gudbjartsson, D., Agnarsson, B.A., Sigurdsson, A., Benediktsdottir, K.R., Blondal, T., Jakobsdottir, M., Stacey, S.N., Kostic, J., Kristinsson, K.T., Birgisdottir, B., Ghosh, S., Magnusdottir, D.N., Thorlacius, S., Thorleifsson, G., Zheng, S.L., Sun, J., Chang, B.L., Elmore, J.B., Breyer, J.P., McReynolds, K.M., Bradley, K.M., Yaspan, B.L., Wiklund, F., Stattin, P., Lindstrom, S., Adami, H.O., McDonnell, S.K., Schaid, D.J., Cunningham, J.M., Wang, L., Cerhan, J.R., St Sauver, J.L., Isaacs, S.D., Wiley, K.E., Partin, A.W., Walsh, P.C., Polo, S., Ruiz-Echarri, M., Navarrete, S., Fuertes, F., Saez, B., Godino, J., Weijerman, P.C., Swinkels, D.W., Aben, K.K., Witjes, J.A., Suarez, B.K., Helfand, B.T., Frigge, M.L., Kristjansson, K., Ober, C., Jonsson, E., Einarsson, G.V., Xu, J., Gronberg, H., Smith, J.R., Thibodeau, S.N., Isaacs, W.B., Catalona, W.J., Mayordomo, J.I., Kiemeney, L.A., Barkardottir, R.B., Gulcher, J.R., Thorsteinsdottir, U., Kong, A. and Stefansson, K. (2008) Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. *Nat Genet,* 40, 281-3.**
10. **Eeles, R.A., Kote-Jarai, Z., Giles, G.G., Olama, A.A., Guy, M., Jugurnauth, S.K., Mulholland, S., Leongamornlert, D.A., Edwards, S.M., Morrison, J., Field, H.I., Southey, M.C., Severi, G., Donovan, J.L., Hamdy, F.C., Dearnaley, D.P., Muir, K.R., Smith, C., Bagnato, M., Ardern-Jones, A.T., Hall, A.L., O'Brien, L.T., Gehr-Swain, B.N., Wilkinson, R.A., Cox, A., Lewis, S., Brown, P.M., Jhavar, S.G., Tymrakiewicz, M., Lophatananon, A., Bryant, S.L., Horwich, A., Huddart, R.A., Khoo, V.S., Parker, C.C., Woodhouse, C.J., Thompson, A., Christmas, T., Ogden, C., Fisher, C., Jamieson, C., Cooper, C.S., English, D.R., Hopper, J.L., Neal, D.E. and Easton, D.F. (2008) Multiple newly identified loci associated with prostate cancer susceptibility. *Nat Genet,* 40, 316-21.**
11. **Yeager, M., Chatterjee, N., Ciampa, J., Jacobs, K.B., Gonzalez-Bosquet, J., Hayes, R.B., Kraft, P., Wacholder, S., Orr, N., Berndt, S., Yu, K., Hutchinson, A., Wang, Z., Amundadottir, L., Feigelson, H.S., Thun, M.J., Diver, W.R., Albanes, D., Virtamo, J., Weinstein, S., Schumacher, F.R., Cancel-Tassin, G., Cussenot, O., Valeri, A., Andriole, G.L., Crawford, E.D., Haiman, C.A., Henderson, B., Kolonel, L., Le Marchand, L., Siddiq, A., Riboli, E., Key, T.J., Kaaks, R., Isaacs, W., Isaacs, S., Wiley, K.E., Gronberg, H., Wiklund, F., Stattin, P., Xu, J., Zheng, S.L., Sun, J., Vatten, L.J., Hveem, K., Kumle, M., Tucker, M., Gerhard, D.S., Hoover, R.N., Fraumeni, J.F., Jr., Hunter, D.J., Thomas, G. and Chanock, S.J. (2009) Identification of a new prostate cancer susceptibility locus on chromosome 8q24. *Nat Genet,* 41, 1055-7.**
12. **Gudmundsson, J., Sulem, P., Gudbjartsson, D.F., Blondal, T., Gylfason, A., Agnarsson, B.A., Benediktsdottir, K.R., Magnusdottir, D.N., Orlygsdottir, G., Jakobsdottir, M., Stacey, S.N., Sigurdsson, A., Wahlfors, T., Tammela, T., Breyer, J.P., McReynolds, K.M., Bradley, K.M., Saez, B., Godino, J., Navarrete, S., Fuertes, F., Murillo, L., Polo, E., Aben, K.K., van Oort, I.M., Suarez, B.K., Helfand, B.T., Kan, D., Zanon, C., Frigge, M.L., Kristjansson, K., Gulcher, J.R., Einarsson, G.V., Jonsson, E., Catalona, W.J., Mayordomo, J.I., Kiemeney, L.A., Smith, J.R., Schleutker, J., Barkardottir, R.B., Kong, A., Thorsteinsdottir, U., Rafnar, T. and Stefansson, K. (2009) Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. *Nat Genet,* 41, 1122-6.**
13. **Eeles, R.A., Kote-Jarai, Z., Al Olama, A.A., Giles, G.G., Guy, M., Severi, G., Muir, K., Hopper, J.L., Henderson, B.E., Haiman, C.A., Schleutker, J., Hamdy, F.C., Neal, D.E., Donovan, J.L.,**

Stanford, J.L., Ostrander, E.A., Ingles, S.A., John, E.M., Thibodeau, S.N., Schaid, D., Park, J.Y., Spurdle, A., Clements, J., Dickinson, J.L., Maier, C., Vogel, W., Dork, T., Rebbeck, T.R., Cooney, K.A., Cannon-Albright, L., Chappuis, P.O., Hutter, P., Zeegers, M., Kaneva, R., Zhang, H.W., Lu, Y.J., Foulkes, W.D., English, D.R., Leongamornlert, D.A., Tymrakiewicz, M., Morrison, J., Ardern-Jones, A.T., Hall, A.L., O'Brien, L.T., Wilkinson, R.A., Saunders, E.J., Page, E.C., Sawyer, E.J., Edwards, S.M., Dearnaley, D.P., Horwich, A., Huddart, R.A., Khoo, V.S., Parker, C.C., Van As, N., Woodhouse, C.J., Thompson, A., Christmas, T., Ogden, C., Cooper, C.S., Southey, M.C., Lophatananon, A., Liu, J.F., Kolonel, L.N., Le Marchand, L., Wahlfors, T., Tammela, T.L., Auvinen, A., Lewis, S.J., Cox, A., FitzGerald, L.M., Koopmeiners, J.S., Karyadi, D.M., Kwon, E.M., Stern, M.C., Corral, R., Joshi, A.D., Shahabi, A., McDonnell, S.K., Sellers, T.A., Pow-Sang, J., Chambers, S., Aitken, J., Gardiner, R.A., Batra, J., Kedda, M.A., Lose, F., Polanowski, A., Patterson, B., Serth, J., Meyer, A., Luedeke, M., Stefflova, K., Ray, A.M., Lange, E.M., Farnham, J., Khan, H., Slavov, C., Mitkova, A., Cao, G. and Easton, D.F. (2009) Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. *Nat Genet*, 41, 1116-21.

14. **Sun, J., Purcell, L., Gao, Z., Isaacs, S.D., Wiley, K.E., Hsu, F.C., Liu, W., Duggan, D., Carpten, J.D., Gronberg, H., Xu, J., Chang, B.L., Partin, A.W., Walsh, P.C., Isaacs, W.B. and Zheng, S.L. (2008) Association between sequence variants at 17q12 and 17q24.3 and prostate cancer risk in European and African Americans. *Prostate*, 68, 691-7.**
15. **Hsu, F.C., Sun, J., Wiklund, F., Isaacs, S.D., Wiley, K.E., Purcell, L.D., Gao, Z., Stattin, P., Zhu, Y., Kim, S.T., Zhang, Z., Liu, W., Chang, B.L., Walsh, P.C., Duggan, D., Carpten, J.D., Isaacs, W.B., Gronberg, H., Xu, J. and Zheng, S.L. (2009) A novel prostate cancer susceptibility locus at 19q13. *Cancer Res*, 69, 2720-3.**
16. **Sun, J., Zheng, S.L., Wiklund, F., Isaacs, S.D., Li, G., Wiley, K.E., Kim, S.T., Zhu, Y., Zhang, Z., Hsu, F.C., Turner, A.R., Stattin, P., Liu, W., Kim, J.W., Duggan, D., Carpten, J., Isaacs, W., Gronberg, H., Xu, J. and Chang, B.L. (2009) Sequence variants at 22q13 are associated with prostate cancer risk. *Cancer Res*, 69, 10-5.**
17. **Zheng, S.L., Stevens, V.L., Wiklund, F., Isaacs, S.D., Sun, J., Smith, S., Pruett, K., Wiley, K.E., Kim, S.T., Zhu, Y., Zhang, Z., Hsu, F.C., Turner, A.R., Johansson, J.E., Liu, W., Kim, J.W., Chang, B.L., Duggan, D., Carpten, J., Rodriguez, C., Isaacs, W., Gronberg, H. and Xu, J. (2009) Two independent prostate cancer risk-associated Loci at 11q13. *Cancer Epidemiol Biomarkers Prev*, 18, 1815-20.**
18. **Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. and Sham, P.C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, 81, 559-75.**
19. **Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*, 38, 904-9.**
20. **Manning, A.K., LaValley, M., Liu, C.T., Rice, K., An, P., Liu, Y., Miljkovic, I., Rasmussen-Torvik, L., Harris, T.B., Province, M.A., Borecki, I.B., Florez, J.C., Meigs, J.B., Cupples, L.A. and Dupuis, J. (2011) Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP x environment regression coefficients. *Genet Epidemiol*, 35, 11-8.**
21. **Agresti, A. (2002) Categorical Data Analysis. *Wiley Series in Probability and Statistics, Second Edition (wiley and Sons)*.**
22. **Cordell, H.J. (2009) Detecting gene-gene interactions that underlie human diseases. *Nat Rev Genet*, 10, 392-404.**
23. **Janssen, J.W., Vaandrager, J.W., Heuser, T., Jauch, A., Kluin, P.M., Geelen, E., Bergsagel, P.L., Kuehl, W.M., Drexler, H.G., Otsuki, T., Bartram, C.R. and Schuuring, E. (2000) Concurrent activation of a novel putative transforming gene, myeov, and cyclin D1 in a subset of multiple myeloma cell lines with t(11;14)(q13;q32). *Blood*, 95, 2691-8.**

24. Janssen, J.W., Imoto, I., Inoue, J., Shimada, Y., Ueda, M., Imamura, M., Bartram, C.R. and Inazawa, J. (2002) MYEOV, a gene at 11q13, is coamplified with CCND1, but epigenetically inactivated in a subset of esophageal squamous cell carcinomas. *J Hum Genet*, 47, 460-4.

25. Janssen, J.W., Cuny, M., Orsetti, B., Rodriguez, C., Valles, H., Bartram, C.R., Schuuring, E. and Theillet, C. (2002) MYEOV: a candidate gene for DNA amplification events occurring centromeric to CCND1 in breast cancer. *Int J Cancer*, 102, 608-14.

26. Specht, K., Haralambieva, E., Bink, K., Kremer, M., Mandl-Weber, S., Koch, I., Tomer, R., Hofler, H., Schuuring, E., Kluin, P.M., Fend, F. and Quintanilla-Martinez, L. (2004) Different mechanisms of cyclin D1 overexpression in multiple myeloma revealed by fluorescence in situ hybridization and quantitative analysis of mRNA levels. *Blood*, 104, 1120-6.

27. Moss, A.C., Lawlor, G., Murray, D., Tighe, D., Madden, S.F., Mulligan, A.M., Keane, C.O., Brady, H.R., Doran, P.P. and MacMathuna, P. (2006) ETV4 and Myeov knockdown impairs colon cancer cell line proliferation and invasion. *Biochem Biophys Res Commun*, 345, 216-21.

28. Hatch, E.M., Kulukian, A., Holland, A.J., Cleveland, D.W. and Stearns, T. (2010) Cep152 interacts with Plk4 and is required for centriole duplication. *J Cell Biol*, 191, 721-9.

29. Kalay, E., Yigit, G., Aslan, Y., Brown, K.E., Pohl, E., Bicknell, L.S., Kayserili, H., Li, Y., Tuysuz, B., Nurnberg, G., Kiess, W., Koegl, M., Baessmann, I., Buruk, K., Toraman, B., Kayipmaz, S., Kul, S., Ikbal, M., Turner, D.J., Taylor, M.S., Aerts, J., Scott, C., Milstein, K., Dollfus, H., Wieczorek, D., Brunner, H.G., Hurles, M., Jackson, A.P., Rauch, A., Nurnberg, P., Karaguzel, A. and Wollnik, B. (2011) CEP152 is a genome maintenance protein disrupted in Seckel syndrome. *Nat Genet*, 43, 23-6.

30. Cui, H., Cruz-Correa, M., Giardiello, F.M., Hutcheon, D.F., Kafonek, D.R., Brandenburg, S., Wu, Y., He, X., Powe, N.R. and Feinberg, A.P. (2003) Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science*, 299, 1753-5.

31. Zhao, R., DeCoteau, J.F., Geyer, C.R., Gao, M., Cui, H. and Casson, A.G. (2009) Loss of imprinting of the insulin-like growth factor II (IGF2) gene in esophageal normal and adenocarcinoma tissues. *Carcinogenesis*, 30, 2117-22.

32. Vorwerk, P., Wex, H., Bessert, C., Hohmann, B., Schmidt, U. and Mittler, U. (2003) Loss of imprinting of IGF-II gene in children with acute lymphoblastic leukemia. *Leuk Res*, 27, 807-12.

33. D F Jarrard, M.J.B., G S Bova and W B Isaacs (1995) Regional loss of imprinting of the insulin-like growth factor II gene occurs in human prostate tissues. *Clin Canc Res*, 1, 8.

34. Veronese, A., Lupini, L., Consiglio, J., Visone, R., Ferracin, M., Fornari, F., Zanesi, N., Alder, H., D'Elia, G., Gramantieri, L., Bolondi, L., Lanza, G., Querzoli, P., Angioni, A., Croce, C.M. and Negrini, M. (2010) Oncogenic role of miR-483-3p at the IGF2/483 locus. *Cancer Res*, 70, 3140-9.

35. Hao, J., Zhang, S., Zhou, Y., Hu, X. and Shao, C. (2011) MicroRNA 483-3p suppresses the expression of DPC4/Smad4 in pancreatic cancer. *FEBS Lett*, 585, 207-13.

36. Mohr, F., Dohner, K., Buske, C. and Rawat, V.P. (2011) TET genes: new players in DNA demethylation and important determinants for stemness. *Exp Hematol*, 39, 272-81.

37. Lorsbach, R.B., Moore, J., Mathew, S., Raimondi, S.C., Mukatira, S.T. and Downing, J.R. (2003) TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia*, 17, 637-41.

38. Abdel-Wahab, O., Mullally, A., Hedvat, C., Garcia-Manero, G., Patel, J., Wadleigh, M., Malinge, S., Yao, J., Kilpivaara, O., Bhat, R., Huberman, K., Thomas, S., Dolgalev, I., Heguy, A., Paietta, E., Le Beau, M.M., Beran, M., Tallman, M.S., Ebert, B.L., Kantarjian, H.M., Stone, R.M., Gilliland, D.G., Crispino, J.D. and Levine, R.L. (2009) Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood*, 114, 144-7.

39. Coopman, P.J. and Mueller, S.C. (2006) The Syk tyrosine kinase: a new negative regulator in tumor growth and progression. *Cancer Lett*, 241, 159-73.

40. Ma, L., Dong, S., Zhang, P., Xu, N., Yan, H., Liu, H., Li, Y. and Zhou, Q. (2010) The relationship between methylation of the Syk gene in the promoter region and the genesis of lung cancer. *Clin Lab*, 56, 407-16.
41. Yuan, Y., Mendez, R., Sahin, A. and Dai, J.L. (2001) Hypermethylation leads to silencing of the SYK gene in human breast cancer. *Cancer Res*, 61, 5558-61.
42. Wang, B. (2011) Cancer Cells Exploit the Eph-Ephrin System to Promote Invasion and Metastasis: Tales of Unwitting Partners. *Sci Signal*, 4.
43. Shan, Z., Parker, T. and Wiest, J.S. (2004) Identifying novel homozygous deletions by microsatellite analysis and characterization of tumor suppressor candidate 1 gene, TUSC1, on chromosome 9p in human lung cancer. *Oncogene*, 23, 6612-20.
44. Wan, X., Yang, C., Yang, Q., Xue, H., Fan, X., Tang, N.L. and Yu, W. (2010) BOOST: A fast approach to detecting gene-gene interactions in genome-wide case-control studies. *Am J Hum Genet*, 87, 325-40.
45. Bell, J.T., Timpson, N.J., Rayner, N.W., Zeggini, E., Frayling, T.M., Hattersley, A.T., Morris, A.P. and McCarthy, M.I. (2011) Genome-wide association scan allowing for epistasis in type 2 diabetes. *Ann Hum Genet*, 75, 10-9.
46. Culverhouse, R., Suarez, B.K., Lin, J. and Reich, T. (2002) A perspective on epistasis: limits of models displaying no main effect. *Am J Hum Genet*, 70, 461-71.
47. Ritchie, M.D., Hahn, L.W., Roodi, N., Bailey, L.R., Dupont, W.D., Parl, F.F. and Moore, J.H. (2001) Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am J Hum Genet*, 69, 138-47.
48. Marchini, J., Donnelly, P. and Cardon, L.R. (2005) Genome-wide strategies for detecting multiple loci that influence complex diseases. *Nat Genet*, 37, 413-7.
49. Wicki, A., Lehembre, F., Wick, N., Hantusch, B., Kerjaschki, D. and Christofori, G. (2006) Tumor invasion in the absence of epithelial-mesenchymal transition: podoplanin-mediated remodeling of the actin cytoskeleton. *Cancer Cell*, 9, 261-72.
50. Navarro, A., Perez, R.E., Rezaiekhaligh, M.H., Mabry, S.M. and Ekekezie, II (2011) Polarized migration of lymphatic endothelial cells is critically dependent on podoplanin regulation of Cdc42. *Am J Physiol Lung Cell Mol Physiol*, 300, L32-42.
51. Kaneko, M., Kato, Y., Kunita, A., Fujita, N., Tsuruo, T. and Osawa, M. (2004) Functional sialylated O-glycan to platelet aggregation on Aggrus (T1alpha/Podoplanin) molecules expressed in Chinese hamster ovary cells. *J Biol Chem*, 279, 38838-43.
52. Schacht, V., Ramirez, M.I., Hong, Y.K., Hirakawa, S., Feng, D., Harvey, N., Williams, M., Dvorak, A.M., Dvorak, H.F., Oliver, G. and Detmar, M. (2003) T1alpha/podoplanin deficiency disrupts normal lymphatic vasculature formation and causes lymphedema. *EMBO J*, 22, 3546-56.
53. Martin-Villar, E., Scholl, F.G., Gamallo, C., Yurrita, M.M., Munoz-Guerra, M., Cruces, J. and Quintanilla, M. (2005) Characterization of human PA2.26 antigen (T1alpha-2, podoplanin), a small membrane mucin induced in oral squamous cell carcinomas. *Int J Cancer*, 113, 899-910.
54. Schacht, V., Dadras, S.S., Johnson, L.A., Jackson, D.G., Hong, Y.K. and Detmar, M. (2005) Up-regulation of the lymphatic marker podoplanin, a mucin-type transmembrane glycoprotein, in human squamous cell carcinomas and germ cell tumors. *Am J Pathol*, 166, 913-21.
55. Rahadiani, N., Ikeda, J., Makino, T., Tian, T., Qiu, Y., Mamat, S., Wang, Y., Doki, Y., Aozasa, K. and Morii, E. (2010) Tumorigenic role of podoplanin in esophageal squamous-cell carcinoma. *Ann Surg Oncol*, 17, 1311-23.
56. Mishima, K., Kato, Y., Kaneko, M.K., Nakazawa, Y., Kunita, A., Fujita, N., Tsuruo, T., Nishikawa, R., Hirose, T. and Matsutani, M. (2006) Podoplanin expression in primary central nervous system germ cell tumors: a useful histological marker for the diagnosis of germinoma. *Acta Neuropathol*, 111, 563-8.
57. Wicki, A. and Christofori, G. (2007) The potential role of podoplanin in tumour invasion. *Br J Cancer*, 96, 1-5.

58. Raica, M., Cimpean, A.M. and Ribatti, D. (2008) The role of podoplanin in tumor progression and metastasis. ***Anticancer Res*****, 28, 2997-3006.**

59. Kanner, W.A., Galgano, M.T. and Atkins, K.A. (2010) Podoplanin expression in basal and myoepithelial cells: utility and potential pitfalls. ***Appl Immunohistochem Mol Morphol*****, 18, 226-30.**

60. Georgopoulos, K., Bigby, M., Wang, J.H., Molnar, A., Wu, P., Winandy, S. and Sharpe, A. (1994) The Ikaros gene is required for the development of all lymphoid lineages. ***Cell*****, 79, 143-56.**

61. Ezzat, S., Zhu, X., Loeper, S., Fischer, S. and Asa, S.L. (2006) Tumor-derived Ikaros 6 acetylates the Bcl-XL promoter to up-regulate a survival signal in pituitary cells. ***Mol Endocrinol*****, 20, 2976-86.**

62. Sun, L., Heerema, N., Crotty, L., Wu, X., Navara, C., Vassilev, A., Sensel, M., Reaman, G.H. and Uckun, F.M. (1999) Expression of dominant-negative and mutant isoforms of the antileukemic transcription factor Ikaros in infant acute lymphoblastic leukemia. ***Proc Natl Acad Sci U S A*****, 96, 680-5.**

63. Ezzat, S., Yu, S. and Asa, S.L. (2003) Ikaros isoforms in human pituitary tumors: distinct localization, histone acetylation, and activation of the 5' fibroblast growth factor receptor-4 promoter. ***Am J Pathol*****, 163, 1177-84.**

64. Yang, L., Luo, Y. and Wei, J. (2010) Integrative genomic analyses on Ikaros and its expression related to solid cancer prognosis. ***Oncol Rep*****, 24, 571-7.**

65. Wang, Q., Li, W., Zhang, Y., Yuan, X., Xu, K., Yu, J., Chen, Z., Beroukhim, R., Wang, H., Lupien, M., Wu, T., Regan, M.M., Meyer, C.A., Carroll, J.S., Manrai, A.K., Janne, O.A., Balk, S.P., Mehra, R., Han, B., Chinnaiyan, A.M., Rubin, M.A., True, L., Fiorentino, M., Fiore, C., Loda, M., Kantoff, P.W., Liu, X.S. and Brown, M. (2009) Androgen receptor regulates a distinct transcription program in androgen-independent prostate cancer. ***Cell*****, 138, 245-56.**

66. Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F., Wang, Q., Bekiranov, S., Sementchenko, V., Fox, E.A., Silver, P.A., Gingeras, T.R., Liu, X.S. and Brown, M. (2006) Genome-wide analysis of estrogen receptor binding sites. ***Nat Genet*****, 38, 1289-97.**

Supplementary Table 1. Clinical and demographic characteristics of subjects in CAPS

Characteristics	# (%) of cases			# (%) of
	Aggressive (N=686)	Localized (N=795)	All cases (N=1483)	controls (N=519)
Age at enrollment (Year)				
Mean (sd)	67.91(7.25)	64.59(6.55)	66.13(7.07)	67.24(7.35)
Age at disgnosis				
≤ 65	60.24(3.76)	59.55(3.66)	59.81(3.71)	NA
> 65	72.5(4.35)	70.5(4.15)	71.6(4.38)	NA
Family History (first-degree relatives)				
No	569(82.94)	623(78.36)	1192(80.38)	466(89.79)
Yes	117(17.06)	172(21.64)	289(19.49)	53(10.21)
Missing data	0(0)	0(0)	2(.13)	0(0)
PSA levels at diagnosis for cases or at enrollment for controls (ng/ml)				
≤ 4	20(2.92)	98(12.33)	118(7.96)	413(79.58)
4-9.99	94(13.70)	430(54.08)	524(35.33)	81(15.61)
10-19.99	117(17.06)	188(23.65)	305(20.57)	20(3.85)
20-49.99	150(21.87)	72(9.06)	222(14.97)	4(.77)
50-99.99	128(18.66)	0(0)	128(8.63)	1(.19)
≥ 100	172(25.07)	0(0)	172(11.60)	0(0)
Missing	5(.73)	7(.88)	14(.94)	0(0)
T-stage				
T0	1(.13)	4(.5)	5(.34)	NA
T1	83(10.44)	470(59.12)	553(37.29)	NA
T2	138(17.36)	316(39.75)	454(30.61)	NA
T3	399(50.19)	0(0)	399(26.90)	NA
T4	58(7.30)	0(0)	58(3.91)	NA
TX	7(.88)	5(.63)	14(.94)	NA
N-stage				
N0	130(16.35)	123(15.47)	253(17.06)	NA
N1	45(5.66)	0(0)	45(3.03)	NA
NX	511(64.28)	672(84.53)	1185(79.90)	NA
M-stage				
M0	324(47.23)	298(37.48)	622(41.94)	NA
M1	159(23.18)	0(0)	159(10.72)	NA
MX	203(29.59)	497(62.52)	702(47.34)	NA
Gleason (biopsy)				
≤ 4	7(1.02)	0(0)	7(.47)	NA
5	25(3.64)	1(.13)	26(1.75)	NA
6	86(12.54)	791(99.50)	877(59.14)	NA
7	218(31.78)	1(.13)	219(14.77)	NA
8	156(22.74)	0(0)	156(10.52)	NA
9	108(15.74)	0(0)	108(7.28)	NA
10	13(1.90)	0(0)	13(.88)	NA
Missing	73(10.64)	2(.25)	77(5.19)	NA

Supplementary Table 2. Clinical and demographic characteristics of subjects in JHH

Characteristics	# (%) of cases All cases (N=1,964)
Age at diagnosis (Year)	
Mean (sd)	57.75 (6.81)
Missing	9
Age at diagnosis (Year)	
≤ 65	1704 (87.16)
> 65	251 (12.84)
PSA levels at diagnosis (ng/ml)	
≤ 4	310 (16.28)
4.01-9.99	1220 (64.08)
10-19.99	256 (13.45)
20-49.99	66 (3.47)
50-99.99	23 (1.21)
≥ 100	29 (1.52)
Missing	60
T-stage	
T2	1247 (63.49)
T3a	454 (23.12)
T3b	105 (5.35)
T3c	12 (0.61)
T3X	7 (0.36)
T4	1 (0.05)
Missing	138 (7.03)
N-stage	
N0	1782 (97.38)
N1	37 (2.02)
N2	1 (0.05)
NX	10 (0.55)
M-stage	
M0	NA
M1	NA
MX	1828
Gleason (biopsy)	
≤ 4	0
5	41 (2.13)
6	1118 (58.17)
7(3+4 or unspecific)	474 (24.66)
7(4+3)	133 (6.92)
8	76 (3.95)
9	74 (3.85)
10	6 (0.31)
Missing	42

Supplementary Table 3. Inflation factor for the meta-analysis for 32 SNPs.

SNP1	N Obs	Median_p	Median_chisq	Inflation Factor
rs10086908	1117528	0.50	0.46	1.02
rs10486567	1117530	0.50	0.45	0.99
rs10896449	1117520	0.50	0.46	1.00
rs10934853	1117529	0.49	0.47	1.03
rs10993994	1117531	0.50	0.46	1.02
rs11649743	1117531	0.50	0.46	1.00
rs12418451	1117531	0.50	0.46	1.00
rs12621278	1117530	0.50	0.45	0.98
rs1447295	1117526	0.50	0.46	1.01
rs1465618	1117531	0.50	0.46	1.01
rs1512268	1117525	0.50	0.45	1.00
rs1571801	1117531	0.50	0.46	1.02
rs16901979	1117504	0.51	0.44	0.97
rs17021918	1117529	0.50	0.46	1.00
rs1859962	1117526	0.50	0.45	0.98
rs2660753	1117527	0.50	0.45	1.00
rs2735839	1117530	0.50	0.46	1.01
rs2928679	1117528	0.50	0.46	1.00
rs4430796	1117531	0.50	0.45	0.98
rs445114	1117531	0.50	0.46	1.01
rs4962416	1117530	0.50	0.47	1.02
rs5759167	1117531	0.50	0.46	1.01
rs5945619	1117527	0.50	0.45	1.00
rs6465657	1117522	0.50	0.46	1.02
rs6983267	1117530	0.50	0.46	1.01
rs7127900	1117531	0.50	0.46	1.01
rs721048	1117531	0.49	0.47	1.03
rs7679673	1117531	0.50	0.46	1.01
rs8102476	1117531	0.50	0.45	0.99
rs887391	1117530	0.50	0.46	1.01
rs9364554	1117531	0.50	0.45	0.99
rs9623117	1117531	0.50	0.45	0.98

Supplementary Table 4. Results for top SNPs that interact with the known PCa-risk SNPs (p<1.0E-05 in meta-analysis)

SNP 1			SNP 2							Meta-results		CAPS		JHH		CGEMS	
CHR	SNP	Gene	CHR	SNP	Position	Minor Allele	Gene	Location	Relative Position	P	OR	P	OR	P	OR	P	OR
2	rs12621278	*ITGA6*	3	rs1002979	113980570	C	*CD200R1L*	Intergenic	36676	3.10E-06	0.63(0.52-0.76)	1.45E-01	0.72(0.46-1.12)	2.66E-04	0.61(0.46-0.79)	9.29E-03	0.62(0.43-0.89)
2	rs12621278		8	rs1402649	20900514	G	*LOC286114*	Intergenic	3604	7.16E-06	1.58(1.29-1.92)	2.13E-01	1.33(0.85-2.10)	8.86E-05	1.74(1.32-2.29)	3.71E-02	1.48(1.02-2.14)
2	rs1465618	*THADA*	4	rs11735008	393303	G	*ABCA11P*	Intergenic	15921	6.65E-06	0.76(0.68-0.86)	5.04E-03	0.69(0.54-0.90)	3.11E-02	0.83(0.71-0.98)	1.35E-03	0.69(0.55-0.86)
2	rs1465618		13	rs9567349	43535405	G	*NCRNA00284*	Intergenic	32806	3.93E-06	0.61(0.49-0.75)	4.57E-04	0.41(0.25-0.67)	3.24E-02	0.72(0.54-0.97)	3.97E-03	0.57(0.39-0.84)
2	rs721048	*EHBP1*	1	rs3820259	239046585	C	*RGS7*	Intron		9.54E-06	1.39(1.20-1.60)	4.40E-02	1.45(1.01-2.08)	9.81E-05	1.46(1.21-1.76)	1.95E-01	1.20(0.91-1.59)
2	rs721048		6	rs9348131	166770788	T	*RPS6KA2*	Intron		6.11E-06	0.70(0.60-0.82)	2.45E-01	0.81(0.57-1.15)	6.83E-04	0.70(0.57-0.86)	3.19E-03	0.63(0.47-0.86)
2	rs721048		7	rs7809487	37110591	A	*ELMO1*	Intron		6.05E-06	0.72(0.62-0.83)	2.65E-01	0.82(0.58-1.16)	1.35E-04	0.69(0.57-0.83)	2.03E-02	0.72(0.55-0.95)
2	rs721048		13	rs9546364	82850742	T	*SLITRK1*	Intergenic	498602	6.33E-06	0.64(0.53-0.78)	6.88E-04	0.48(0.31-0.73)	2.75E-05	0.56(0.43-0.74)	9.45E-01	1.01(0.70-1.46)
3	rs10934853	*EEFSEC*	9	rs7847271	116870633	A	*TNC*	Intron		3.85E-06	0.67(0.56-0.79)	7.73E-03	0.60(0.41-0.87)	2.56E-03	0.68(0.54-0.88)	1.87E-02	0.69(0.50-0.94)
3	rs10934853		14	rs12433148	45666182	A	*RPL10L*	Intergenic	523788	9.25E-06	1.33(1.17-1.50)	6.91E-03	1.48(1.11-1.96)	4.50E-05	1.43(1.20-1.69)	6.32E-01	1.06(0.83-1.35)
3	rs10934853		14	rs2400997	100796860	T	*MIR656*	Intergenic	193969	2.51E-06	1.25(1.14-1.38)	4.80E-05	1.57(1.26-1.95)	6.76E-02	1.13(0.99-1.28)	1.95E-03	1.34(1.11-1.60)
3	rs10934853		18	rs998124	40979660	G	*MIR4319*	Intergenic	-175531	5.21E-06	1.33(1.18-1.51)	3.56E-02	1.39(1.02-1.88)	3.64E-03	1.28(1.08-1.51)	3.56E-03	1.42(1.12-1.80)
3	rs2660753	*NA*	5	rs7717572	66869013	A	*CD180*	Intergenic	-340640	3.39E-06	1.94(1.47-2.56)	2.84E-01	1.60(0.68-3.78)	2.80E-05	2.12(1.49-3.02)	5.80E-02	1.69(0.98-2.89)
3	rs2660753		6	rs319097	107852552	C	*PDSS2*	Intron		9.59E-06	1.35(1.18-1.54)	2.02E-01	1.27(0.88-1.85)	3.15E-06	1.49(1.26-1.76)	5.53E-01	1.08(0.83-1.41)
3	rs2660753		13	rs7139820	106284593	A	*ARGLU1*	Intergenic	-266078	5.64E-06	0.56(0.44-0.72)	1.47E-02	0.46(0.25-0.86)	2.07E-05	0.49(0.35-0.68)	4.97E-01	0.85(0.53-1.36)
4	rs17021918	*PDLIM5*	3	rs9757252	86977168	T	*VGLL3*	Intergenic	92645	4.73E-06	1.25(1.13-1.37)	8.45E-03	1.35(1.08-1.69)	2.23E-03	1.22(1.07-1.38)	2.16E-02	1.24(1.03-1.50)
4	rs17021918		8	rs2921007	8269681	A	*SGK223*	Intron		6.10E-06	1.40(1.21-1.63)	1.79E-02	1.60(1.08-2.35)	3.59E-06	1.58(1.30-1.92)	8.40E-01	1.03(0.78-1.36)
4	rs7679673	*TET2*	9	rs290258	92595560	G	*SYK*	Intergenic	-8273	1.49E-06	0.75(0.67-0.84)	2.11E-03	0.66(0.51-0.86)	3.01E-03	0.78(0.67-0.92)	1.39E-02	0.75(0.59-0.94)
4	rs7679673		11	rs11605083	15311822	C	*INSC*	Intergenic	86492	4.42E-06	1.28(1.15-1.43)	3.28E-01	1.13(0.89-1.44)	8.20E-04	1.28(1.11-1.48)	9.71E-04	1.42(1.15-1.75)
4	rs7679673		22	rs5751168	21175240	T	*ZNF280B*	Intron		4.11E-06	1.44(1.23-1.67)	4.75E-05	2.19(1.50-3.19)	3.38E-02	1.25(1.02-1.53)	9.09E-03	1.48(1.10-1.99)
6	rs9364554	*NA*	6	rs9351730	69351206	A	*BAI3*	Intergenic	-51147	4.98E-06	1.25(1.13-1.37)	6.97E-02	1.22(0.98-1.52)	1.01E-03	1.25(1.09-1.42)	8.99E-03	1.27(1.06-1.52)
7	rs10486567	*JAZF1*	3	rs1795355	41574530	T	*ULK4*	Intron		9.46E-06	0.79(0.71-0.88)	3.37E-02	0.77(0.60-0.98)	9.11E-03	0.83(0.72-0.95)	2.14E-03	0.73(0.60-0.89)
7	rs10486567		3	rs11720607	174325971	G	*SPATA16*	Intron		4.87E-06	0.73(0.63-0.83)	2.45E-03	0.62(0.45-0.84)	2.34E-03	0.75(0.62-0.90)	5.43E-02	0.77(0.58-1.00)
7	rs6465657	*LMTK2*	3	rs12485321	124986	A	*CHL1*	Intergenic	-88664	2.66E-06	0.81(0.74-0.88)	7.36E-02	0.83(0.67-1.02)	2.89E-05	0.78(0.69-0.88)	1.09E-01	0.87(0.74-1.03)
7	rs6465657		3	rs6548941	66555095	T				7.23E-06	1.32(1.17-1.49)	8.73E-03	1.49(1.11-2.01)	2.60E-05	1.42(1.21-1.67)	6.02E-01	1.06(0.84-1.34)
7	rs6465657		16	rs8057939	47951777	C	*C16orf78*	Intergenic	-13532	4.71E-06	1.37(1.20-1.57)	3.31E-02	1.43(1.03-1.98)	1.03E-03	1.36(1.13-1.63)	1.70E-02	1.36(1.06-1.75)

Supplementary Table 4 cont'd

SNP 1			SNP 2							Meta-results		CAPS		JHH		CGEMS	
CHR	SNP	Gene	CHR	SNP	Position	Minor Allele	Gene	Location	Relative Position	P	OR	P	OR	P	OR	P	OR
8	rs10086908	*NA*	4	rs7694725	114220405	T	*ANK2*	Intron		1.55E-06	1.36(1.20-1.55)	7.99E-03	1.46(1.10-1.92)	1.41E-05	1.47(1.24-1.75)	3.89E-01	1.11(0.87-1.42)
8	rs10086908		6	rs10456809	17921804	T	*KIF13A*	Intron		4.83E-06	1.25(1.14-1.38)	1.30E-02	1.31(1.06-1.62)	2.28E-03	1.23(1.08-1.40)	1.69E-02	1.26(1.04-1.52)
8	rs10086908		10	rs4917911	102549411	G	*PAX2*	Intron		7.70E-06	1.43(1.22-1.67)	3.40E-02	1.46(1.03-2.07)	2.63E-04	1.48(1.20-1.83)	1.03E-01	1.30(0.95-1.79)
8	rs1447295	*NA*	7	rs7789197	40931652	A	*INHBA*	Intergenic	763474	3.36E-06	0.66(0.56-0.79)	2.57E-03	0.55(0.38-0.81)	7.27E-03	0.72(0.56-0.91)	1.06E-02	0.66(0.48-0.91)
8	rs1447295		9	rs12682851	8002418	G	*C9orf123*	Intergenic	-212619	1.53E-06	0.72(0.63-0.82)	9.66E-03	0.67(0.50-0.91)	2.15E-03	0.75(0.62-0.90)	6.56E-03	0.70(0.54-0.90)
8	rs1447295		10	rs10885582	116317540	T	*ABLIM1*	Intron		3.70E-06	0.73(0.63-0.83)	9.33E-05	0.54(0.39-0.73)	1.27E-02	0.79(0.66-0.95)	3.46E-02	0.75(0.57-0.98)
8	rs1447295		15	rs11637980	94803657	G	*NR2F2*	Intergenic	119161	1.55E-06	0.68(0.58-0.79)	1.67E-03	0.58(0.41-0.82)	5.73E-02	0.81(0.65-1.01)	8.29E-05	0.52(0.38-0.72)
8	rs1512268	*NKX3.1*	6	rs2523395	29810489	A	*LOC285830*	Intron		1.53E-06	1.24(1.14-1.35)	1.89E-01	1.14(0.94-1.39)	4.54E-02	1.13(1.00-1.26)	2.98E-08	1.65(1.38-1.97)
8	rs1512268		7	rs517761	103156254	T	*RELN*	Intron		6.93E-06	0.82(0.75-0.90)	1.46E-01	0.86(0.70-1.05)	2.46E-05	0.78(0.70-0.88)	1.64E-01	0.89(0.75-1.05)
8	rs1512268		13	rs16944141	89663547	A	*MIR622*	Intergenic	-17890	2.98E-06	0.65(0.54-0.78)	1.54E-01	0.76(0.51-1.11)	7.73E-07	0.54(0.42-0.69)	4.43E-01	0.87(0.60-1.25)
8	rs16901979	*NA*	12	rs12317459	81688687	A	*TMTC2*	Intron		3.80E-06	2.06(1.52-2.80)	7.13E-02	1.92(0.94-3.91)	1.71E-04	2.21(1.46-3.33)	4.00E-02	1.88(1.03-3.42)
8	rs2928679	*NKX3.1*	2	rs17198717	181943741	C	*ITGA4*	Intergenic	-86123	7.80E-06	0.80(0.72-0.88)	3.19E-01	0.89(0.70-1.12)	2.91E-05	0.75(0.65-0.86)	6.73E-02	0.84(0.69-1.01)
8	rs445114	*NA*	22	rs6005451	26182183	C	*MN1*	Intergenic	292082	4.10E-06	1.42(1.22-1.65)	4.95E-01	1.13(0.79-1.63)	1.71E-04	1.46(1.20-1.77)	3.28E-03	1.58(1.16-2.14)
8	rs6983267	*NA*	6	rs1011119	19972144	G	*ID4*	Intergenic	23250	7.20E-06	0.81(0.74-0.89)	1.45E-02	0.76(0.61-0.95)	2.85E-03	0.83(0.74-0.94)	1.58E-02	0.80(0.67-0.96)
8	rs6983267		15	rs543686	32855601	T	*ACTC1*	Intergenic	11988	4.02E-06	1.24(1.13-1.35)	1.37E-01	1.18(0.95-1.45)	1.50E-04	1.26(1.12-1.42)	2.52E-02	1.22(1.03-1.46)
9	rs1571801	*DAB2IC*	8	rs2219968	79119213	A	*PKIA*	Intergenic	-471678	6.07E-07	1.30(1.17-1.43)	1.18E-02	1.35(1.07-1.70)	1.64E-04	1.30(1.14-1.50)	3.17E-02	1.24(1.02-1.52)
9	rs1571801		8	rs13264970	83236384	C	*SNX16*	Intergenic	-319308	3.53E-06	0.77(0.69-0.86)	1.50E-02	0.74(0.59-0.94)	5.02E-03	0.80(0.68-0.93)	3.92E-03	0.72(0.58-0.90)
9	rs1571801		10	rs1547851	92364806	T	*HTR7*	Intergenic	125750	7.45E-06	1.59(1.30-1.95)	2.72E-03	1.98(1.27-3.09)	7.35E-03	1.49(1.11-1.99)	2.42E-02	1.52(1.06-2.20)
9	rs1571801		21	rs11702869	19512402	A	*TMPRSS15*	Intergenic	-814561	8.54E-06	0.79(0.71-0.87)	1.11E-01	0.83(0.66-1.04)	4.50E-05	0.74(0.64-0.85)	1.44E-01	0.86(0.70-1.05)
10	rs10993994	*MSMB*	3	rs6766510	12526807	C	*TSEN2*	Intron		1.77E-06	1.58(1.31-1.91)	9.37E-02	1.46(0.94-2.26)	2.98E-05	1.75(1.34-2.27)	5.32E-02	1.41(1.00-1.98)
10	rs10993994		4	rs567404	16810346	G	*QDPR*	Intergenic	286768	5.91E-06	0.81(0.74-0.89)	7.30E-02	0.82(0.66-1.02)	1.16E-04	0.78(0.69-0.89)	7.90E-02	0.85(0.71-1.02)
10	rs4962416	*CTBP2*	5	rs10940579	57166575	C	*ACTBL2*	Intergenic	-352182	3.81E-06	1.32(1.18-1.49)	4.85E-02	1.36(1.00-1.84)	1.77E-03	1.29(1.10-1.51)	4.74E-03	1.39(1.10-1.74)
10	rs4962416		7	rs9649213	97859147	G	*BAIAP2L1*	Intron		1.42E-06	1.28(1.16-1.42)	7.77E-01	1.04(0.80-1.34)	6.18E-06	1.36(1.19-1.55)	1.40E-02	1.27(1.05-1.54)
10	rs4962416		9	rs10810120	14234376	C	*NFIB*	Intron		8.44E-06	1.40(1.21-1.62)	6.39E-02	1.39(0.98-1.96)	5.58E-03	1.32(1.08-1.61)	1.76E-03	1.59(1.19-2.12)
11	rs10896449	*MYEOV*	2	rs13398206	198877341	C	*PLCL1*	Intergenic	154488	3.67E-06	1.24(1.13-1.36)	6.86E-01	1.05(0.84-1.30)	1.29E-04	1.27(1.12-1.44)	1.91E-03	1.33(1.11-1.58)
11	rs10896449		7	rs6968681	130286690	T	*FLJ43663*	Intron		7.31E-06	0.79(0.72-0.88)	5.86E-01	0.94(0.76-1.17)	4.78E-05	0.75(0.65-0.86)	1.14E-02	0.78(0.64-0.95)
11	rs10896449		12	rs17354197	88169501	T	*DUSP6*	Intergenic	96467	8.82E-06	1.41(1.21-1.64)	3.97E-02	1.49(1.02-2.18)	2.35E-03	1.37(1.12-1.67)	1.10E-02	1.45(1.09-1.92)
11	rs10896449		21	rs447988	39410617	T	*PSMG1*	Intergenic	58637	5.79E-06	0.67(0.56-0.80)	1.96E-01	0.76(0.50-1.15)	1.93E-02	0.76(0.60-0.96)	1.73E-05	0.49(0.35-0.68)

Supplementary Table 4 cont'd

SNP 1			SNP 2							Meta-results		CAPS		JHH		CGEMS	
CHR	SNP	Gene	CHR	SNP	Position	Minor Allele	Gene	Location	Relative Position	P	OR	P	OR	P	OR	P	OR
11	rs12418451	*MYEOV*	3	rs1916284	57369806	C	*DNAH12*	Intron		1.31E-06	0.79(0.71-0.87)	1.12E-01	0.83(0.65-1.05)	2.69E-05	0.76(0.66-0.86)	4.50E-02	0.83(0.68-1.00)
11	rs12418451		3	rs10513723	176062702	A	*NAALADL2*	Intron		5.61E-06	1.41(1.21-1.63)	7.22E-03	1.58(1.13-2.21)	1.61E-02	1.27(1.05-1.54)	1.46E-03	1.64(1.21-2.22)
11	rs12418451		8	rs7829048	4689690	C	*CSMD1*	Intron		9.76E-06	0.74(0.65-0.85)	1.22E-02	0.68(0.50-0.92)	4.52E-02	0.84(0.71-1.00)	1.81E-04	0.60(0.46-0.78)
11	rs12418451		14	rs1243647	20094459	A	*RNASE9*	missense	`CG)204P(CCG)	1.42E-06	0.75(0.67-0.84)	6.63E-02	0.77(0.58-1.02)	5.02E-04	0.76(0.65-0.88)	4.91E-03	0.72(0.58-0.91)
11	rs12418451		15	rs784411	46827089	C	*CEP152*	Intron		1.15E-07	1.42(1.25-1.61)	8.83E-03	1.55(1.12-2.16)	5.28E-04	1.34(1.14-1.58)	1.32E-03	1.53(1.18-1.99)
11	rs7127900	*IGF2, IGF2AS*	2	rs3789080	111514002	C	*ACOXL*	Intron		4.37E-06	0.71(0.61-0.82)	2.23E-02	0.69(0.51-0.95)	1.24E-04	0.67(0.54-0.82)	1.32E-01	0.80(0.61-1.07)
11	rs7127900	*INS,TH*	8	rs13258681	124783903	C	*ANXA13*	Intron		3.65E-06	1.32(1.17-1.48)	4.58E-02	1.32(1.01-1.73)	5.43E-04	1.33(1.13-1.56)	1.90E-02	1.31(1.05-1.64)
11	rs7127900		13	rs9594759	41930593	C	*TNFSF11*	Intergenic	-104279	7.96E-06	0.78(0.70-0.87)	3.34E-01	0.88(0.69-1.14)	4.58E-04	0.77(0.67-0.89)	4.43E-03	0.73(0.58-0.91)
11	rs7127900		22	rs12628051	38984222	C	*TNRC6B*	Intron		3.39E-06	1.30(1.17-1.46)	1.82E-03	1.50(1.16-1.93)	6.14E-03	1.24(1.06-1.44)	1.44E-02	1.32(1.06-1.65)
11	rs7127900		22	rs4821941	39005037	G	*TNRC6B*	Intron		4.35E-06	1.30(1.16-1.46)	2.71E-03	1.48(1.15-1.91)	6.61E-03	1.23(1.06-1.44)	1.27E-02	1.33(1.06-1.66)
17	rs11649743	*HNF1B*	6	rs13192613	123324640	T	*CLVS2*	Intergenic	-34641	3.07E-06	1.37(1.20-1.56)	6.67E-03	1.56(1.13-2.16)	3.69E-04	1.38(1.16-1.65)	8.49E-02	1.24(0.97-1.59)
17	rs4430796	*HNF1B*	1	rs731174	37969428	C	*EPHA10*	Intron		4.55E-06	1.27(1.15-1.40)	5.03E-02	1.24(1.00-1.54)	3.20E-02	1.18(1.01-1.38)	1.13E-04	1.41(1.19-1.68)
17	rs4430796		2	rs12694942	158518681	T	*UPP2*	Intergenic	-41256	7.84E-06	0.80(0.73-0.88)	5.22E-03	0.75(0.61-0.92)	1.87E-03	0.79(0.68-0.92)	6.36E-02	0.85(0.72-1.01)
17	rs4430796		3	rs13067734	143445705	C	*GK5*	Intergenic	-18566	8.07E-06	0.80(0.73-0.88)	3.87E-01	0.91(0.74-1.13)	8.47E-02	0.88(0.77-1.02)	2.79E-07	0.63(0.53-0.75)
17	rs4430796		9	rs10812303	25712117	T	*TUSC1*	Intergenic	-43261	5.59E-06	1.34(1.18-1.52)	6.63E-03	1.46(1.11-1.91)	7.57E-03	1.27(1.07-1.52)	9.01E-03	1.37(1.08-1.73)
17	rs4430796		9	rs7855134	70264671	T	*PGM5*	Intron		7.52E-06	1.76(1.38-2.26)	3.79E-02	1.89(1.04-3.44)	1.92E-04	1.93(1.37-2.73)	9.14E-02	1.46(0.94-2.28)
17	rs4430796		12	rs4489787	47097367	C	*ANP32D*	Intergenic	-55348	1.26E-06	0.68(0.58-0.80)	5.26E-01	0.89(0.63-1.27)	1.41E-03	0.70(0.56-0.87)	3.05E-05	0.55(0.42-0.73)
17	rs1859962	*NA*	2	rs16867225	180749531	A	*CWC22*	Intergenic	-169506	3.12E-06	0.64(0.53-0.77)	1.33E-03	0.48(0.30-0.75)	4.80E-03	0.70(0.54-0.90)	1.84E-02	0.65(0.45-0.93)
17	rs1859962		7	rs10277209	108790810	C	*C7orf66*	Intergenic	-478937	3.81E-06	1.36(1.19-1.55)	4.08E-02	1.39(1.01-1.89)	4.04E-03	1.29(1.08-1.53)	1.55E-03	1.52(1.17-1.97)
19	rs2735839	*KLK3*	20	rs6089829	61139481	A	*LOC63930*	Intron		3.21E-06	0.74(0.65-0.84)	2.53E-02	0.71(0.52-0.96)	1.26E-03	0.75(0.64-0.90)	1.10E-02	0.72(0.56-0.93)
19	rs8102476	*NA*	1	rs1866967	29958249	G	*PTPRU*	Intergenic	432337	5.16E-06	0.82(0.75-0.89)	2.70E-02	0.79(0.64-0.97)	5.10E-03	0.85(0.75-0.95)	2.97E-03	0.78(0.66-0.92)
19	rs8102476		10	rs10795917	12091822	G	*UPF2*	Intron		6.79E-07	1.24(1.14-1.36)	1.88E-01	1.15(0.93-1.41)	1.13E-04	1.26(1.12-1.42)	3.14E-03	1.27(1.08-1.50)
19	rs887391	*PPP1R14A*	4	rs735172	5809770	C	*EVC*	Intron		2.03E-06	1.31(1.17-1.46)	6.58E-03	1.43(1.10-1.85)	2.37E-03	1.27(1.09-1.47)	1.02E-02	1.32(1.07-1.63)
19	rs887391		5	rs4463179	13558432	A	*DNAH5*	Intergenic	185005	2.22E-06	0.64(0.53-0.77)	3.00E-02	0.64(0.42-0.96)	8.35E-05	0.59(0.46-0.77)	8.28E-02	0.73(0.51-1.04)
19	rs887391		8	rs2981156	39988790	C	*IDO2*	Intron		8.07E-06	1.31(1.16-1.47)	2.22E-03	1.53(1.17-2.02)	1.06E-03	1.31(1.11-1.53)	1.58E-01	1.18(0.94-1.47)
19	rs887391		12	rs10844540	33349548	A	*SYT10*	Intergenic	70067	7.59E-06	1.40(1.21-1.62)	3.86E-02	1.43(1.02-2.02)	8.40E-02	1.19(0.98-1.45)	6.16E-06	1.96(1.46-2.62)
22	rs5759167	*TTLL1,BIK,*	7	rs12111744	20988023	A	*RPL23P8*	Intergenic	154059	5.95E-06	1.30(1.16-1.46)	3.33E-01	1.14(0.88-1.48)	5.44E-04	1.31(1.12-1.53)	2.50E-03	1.42(1.13-1.78)
22	rs5759167	*MCAT,PACSIN2*	12	rs2711721	45658537	T	*AMIGO2*	Intergenic	97220	2.11E-06	1.28(1.16-1.42)	1.90E-01	1.22(0.91-1.65)	1.75E-04	1.29(1.13-1.47)	9.06E-03	1.30(1.07-1.58)
22	rs5945619	*NUDT10*	7	rs7792744	97325907	C	*ASNS*	Intron		5.62E-06	0.80(0.73-0.88)	9.09E-05	0.74(0.64-0.86)	1.00E+00	1.00(1.00-1.00)	7.82E-03	0.85(0.75-0.96)
22	rs5945619		9	rs1044214	85465379	A	*UBQLN1*	utr3		7.97E-06	1.26(1.14-1.39)	5.88E-04	1.32(1.13-1.55)	1.00E+00	1.00(1.00-1.00)	3.05E-03	1.22(1.07-1.39)
22	rs5945619		18	rs6507016	29181773	T	*C18orf34*	Intron		4.33E-06	0.78(0.71-0.87)	2.54E-02	0.83(0.70-0.98)	1.00E+00	1.00(1.00-1.00)	4.10E-05	0.76(0.66-0.86)
22	rs9623117	*TNRC6B*	3	rs6763848	1487587	A	*CNTN6*	Intergenic	67309	3.90E-06	1.30(1.16-1.45)	7.99E-01	1.03(0.80-1.34)	1.27E-06	1.44(1.24-1.66)	8.57E-02	1.21(0.97-1.51)
22	rs9623117		4	rs1713511	43472127	A	*KCTD8*	Intergenic	398550	7.87E-06	1.31(1.16-1.47)	4.94E-03	1.54(1.14-2.09)	6.71E-03	1.23(1.06-1.44)	1.06E-02	1.36(1.07-1.72)
22	rs9623117		6	rs2844806	30041418	T	*HCG9*	Intergenic	-9453	8.12E-06	1.27(1.14-1.41)	1.82E-01	1.20(0.92-1.55)	3.45E-03	1.23(1.07-1.41)	7.92E-04	1.43(1.16-1.76)
22	rs9623117		6	rs1200562	70960267	C	*COL19A1*	Intron		9.96E-06	0.64(0.53-0.78)	1.33E-01	0.71(0.46-1.11)	8.29E-06	0.54(0.42-0.71)	3.62E-01	0.83(0.57-1.23)

Abbreviations: SNP1 indicate the 32 known PCa-risk SNPs. SNP 2 indicates the interacting SNPs; Chr, chromosome; Relative position is the distance of SNP2 relative to the nearest gene if SNP2 is located in the intergenic region; OR : Odds Ratio; P and OR are for the multiplicative interaction term. P and OR for the Meta-analysis are calculated based on a Cochran-Mantel-Haenszel test. CAPS: PCa case-control study from Sweden; JHH: Johns Hopkins Hospital; CGEMS: the Cancer Genetic Markers of Susceptibility;